

Smoothed Semiparametric Estimation on Multivariate Long Memory Processes

Guilherme Pumi and Sílvia R.C. Lopes

Mathematics Institute
Federal University of Rio Grande do Sul

This version: 13/10/2011

Abstract

In this paper we propose and study a general class of Gaussian Semiparametric Estimators (GSE) of the fractional differencing parameter in the context of long-range dependent multivariate time series. We establish large sample properties of the estimator without assuming Gaussianity. The class of models considered here satisfies simple conditions on the spectral density function, restricted to a small neighborhood of the zero frequency and includes important class of VARFIMA processes. We also present a simulation study to assess the finite sample properties of the proposed estimator based on a smoothed version of the GSE which supports its competitiveness.

Keywords: Fractional integration; Long-range dependence; Semiparametric estimation; Smoothed periodogram; VARFIMA processes.

Mathematical Subject Classification (2010). Primary 62H12, 62F12, 62M10, 60G10, 62M15;

1 Introduction

Let $\mathbf{d} = (d_1, \dots, d_q)' \in (-1/2, 1/2)^q$ and let \mathcal{B} be the shift operator. Consider the q -dimensional weakly stationary process $\{\mathbf{X}_t\}_{t=0}^\infty$ obtained as a stationary solution of the difference equations

$$\text{diag}_{k \in \{1, \dots, q\}} \{(1 - \mathcal{B})^{d_k}\} (\mathbf{X}_t - \mathbb{E}(\mathbf{X}_t)) = \mathbf{Y}_t, \quad (1.1)$$

where $\{\mathbf{Y}_t\}_{t=0}^\infty$ is a q -dimensional weakly stationary process whose spectral density function $f_{\mathbf{Y}}$ is bounded and bounded away from zero. Each coordinate process in (1.1) exhibits long-range dependence whenever the respective parameter $d_i > 0$, in the sense that the spectral density function satisfies $f(\lambda) \sim K\lambda^{-2d_i}$, as $\lambda \rightarrow 0^+$, for some constant $K > 0$ and $i \in \{1, \dots, q\}$.

Processes of the form (1.1) constitute the so-called fractionally integrated processes. As a particular case, consider the situation where the i -th coordinate process $\{Y_t^{(i)}\}_{t=0}^\infty$ follows an ARMA model. In this case, the associated coordinate process $\{X_t^{(i)}\}_{t=0}^\infty$ will be a classical ARFIMA process with the same AR and MA orders and differencing parameter d_i . If the process $\{\mathbf{Y}_t\}_{t=0}^\infty$ is a vectorial ARMA process, then the resulting multivariate process will be the so-called VARFIMA process with differencing parameter $\mathbf{d} = (d_1, \dots, d_q)'$. VARFIMA and, more generally, fractionally integrated processes, are widely used to model multivariate processes with long-range dependence. See, for instance, the recent work of Chiriac and Voev (2011) on modeling and forecasting high frequency data by using VARFIMA and fractionally integrated processes.

The parameter \mathbf{d} in (1.1) determines the spectral density function behavior at the zero frequency as well as the long run autocovariance/autocorrelation structure. Hence, estimation becomes an important matter whenever the long run structure of the process is of interest.

Estimation of the parameter \mathbf{d} in the multivariate case has seen a growing interest in the last years. A maximum likelihood approach was first considered in Sowell (1989), but the computational cost of the author's method is very high. A few years later, Luceño (1996) presented

a computationally cheaper alternative for the maximum likelihood approach based on rewriting and approximating the quadratic form of the Gaussian likelihood function. In a recent work, Tsay (2010) proposed an even faster approach to calculate the exact conditional likelihood based on the multivariate Durbin-Levinson algorithm. Although the maximum likelihood approach usually provides good results, it is still a computationally expensive method.

The works of Fox and Taqqu (1986), Giraitis and Surgailis (1990), among others, provided a rigorous asymptotic theory for (univariate) Gaussian parametric estimates which includes, for instance, $n^{1/2}$ -consistency and asymptotic normality. One drawback is the crucial role played by the Gaussianity assumption in the theory, which also requires strong distributional and regularity conditions and is non-robust with respect to the parametric specification of the model, leading to inconsistent estimates under misspecification.

In the univariate case, Gaussian Semiparametric Estimation (GSE) was first introduced in Künsch (1987) and later rigorously developed by Robinson (1995b). It provides a more robust alternative compared to the parametric one, requiring less distributional assumptions and being more efficient. In the multivariate case, Robinson (1995a) was the first to study and develop a rigorous treatment of a semiparametric estimator. A two-step multivariate GSE has been studied in the work of Lobato (1999), which showed its asymptotic normality under mild conditions, but without relying on Gaussianity. A few years later, Shimotsu (2007) introduced a refinement of Lobato's two-step GSE, which is consistent and asymptotically normal under very mild conditions (Gaussianity is, again, nowhere assumed), but with smaller asymptotic variance than Lobato's estimator. The technique applied in Shimotsu (2007) was a multivariate extension of that in Robinson (1995b), powerful enough to show not only the consistency of the proposed estimator, but also the consistency of Lobato's two-step GSE. Recently, Nielsen (2011) extended the work of Shimotsu (2007) to include the non-stationary case by using the so-called extended periodogram.

The estimator introduced in Shimotsu (2007) is based on the specification of the spectral density function in a neighborhood of the zero frequency. Estimation of the differencing parameter d is obtained through minimization of an objective function, which is derived from the expression of the Gaussian log-likelihood function near the zero frequency. To obtain the objective function, the spectral density is estimated by the periodogram of the process. Although asymptotically unbiased, it is well known that the periodogram is not a consistent estimator of the spectral density and do not even converge to a random variable at all, cf. Grenander (1951). Some authors actually consider the periodogram “*an extremely poor (if not useless) estimate of the spectral density function*” (Priestley, 1981, page 420).

Our contribution to the theory of GSE is two-folded. First, being consistency a highly desirable property of an estimator, we study the consequences of substituting the periodogram in Shimotsu (2007)'s objective function by an arbitrary consistent estimator of the spectral density function. We prove the consistency of the proposed estimator under the same assumptions as in Shimotsu (2007) and no assumption on the spectral density estimator other than consistency. Second, considering Shimotsu (2007)'s objective function with the periodogram substituted by an arbitrary spectral density estimator, we derive necessary conditions under which GSE is consistent and satisfy a multivariate CLT. Gaussianity is nowhere assumed. In order to assess the finite sample properties of the estimators studied here and its competitiveness, we present a simulation study based on simulated VARFIMA process. We apply the so-called smoothed periodogram as an estimator of the spectral density function.

The paper is organized as follows. In the next section, we present some preliminaries concepts and results necessary for this work. In Section 3 we introduce the Smoothed GSE as well as a general class of estimators based on appropriate modifications of the GSE's objective function. Section 4 is devoted to derive the consistency of the proposed estimator while in Section 5, its asymptotic normality is studied. In Section 6 we present some Monte Carlo simulation results to

assess the finite sample performance of the proposed estimator. Conclusions and final remarks are reserved to Section 7. Proofs are presented in Appendix A.

2 Preliminaries

Let $\{\mathbf{X}_t\}_{t=0}^\infty$ be a q -dimensional process specified by (1.1) and assume that the spectral density matrix of \mathbf{Y}_t satisfies $f_{\mathbf{Y}} \sim G$ for a real, symmetric, finite and positive definite matrix G . Let f be the spectral density matrix function of \mathbf{X}_t , so that

$$\mathbb{E}[(\mathbf{X}_t - \mathbb{E}(\mathbf{X}_t))(\mathbf{X}_{t+h} - \mathbb{E}(\mathbf{X}_t))'] = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda,$$

for $h \in \mathbb{N}^* := \mathbb{N} \setminus \{0\}$. Following the reasoning in Shimotsu (2007), the spectral density matrix of \mathbf{X}_t at the Fourier frequencies $\lambda_j = 2\pi j/n$, with $j = 1, \dots, m$ and $m = o(n)$, can be written as

$$f(\lambda_j) \sim \Lambda_j(\mathbf{d}) G \overline{\Lambda_j(\mathbf{d})}' , \quad \text{for } \Lambda_j(\mathbf{d}) = \text{diag}_{k \in \{1, \dots, q\}} \{\Lambda_j^{(k)}(\mathbf{d})\} \quad \text{and} \quad \Lambda_j^{(k)}(\mathbf{d}) = \lambda_j^{-d_k} e^{i(\pi - \lambda_j)d_k/2}, \quad (2.1)$$

where, for a complex matrix A , \overline{A}' denotes the conjugate transpose of A . Let

$$w_n(\lambda) := \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n \mathbf{X}_t e^{it\lambda} \quad \text{and} \quad I_n(\lambda) := w_n(\lambda) \overline{w_n(\lambda)}'$$

be the discrete Fourier transform and the periodogram of \mathbf{X}_t at λ , respectively. In order to define the smoothed periodogram of \mathbf{X}_t , let $W_n(k) := (W_n^{ij}(k))_{i,j=1}^q$ be an array of functions (called weights) and $\{\ell(k)\}_{k=0}^\infty$ be an increasing sequence of positive integers satisfying the following conditions

- A1.** $1/\ell(n) + \ell(n)/n \rightarrow 0$, as n tends to infinity;
- A2.** $W_n^{ij}(k) = W_n^{ij}(-k)$ and $W_n^{ij}(k) \geq 0$, for all k ;
- A3.** $\sum_{|k| \leq \ell(n)} W_n^{ij}(k) = 1$;
- A4.** $\sum_{|k| \leq \ell(n)} W_n^{ij}(k)^2 \rightarrow 0$, as n tends to infinity.

Conditions **A1** - **A4** are standard in the literature of smoothed periodogram. See for instance Priestley (1981). For a Fourier frequency λ_j , let

$$\hat{f}_n(\lambda_j) := \sum_{|k| \leq \ell(n)} W_n(k) \odot w_n(\lambda_{j+k}) \overline{w_n(\lambda_{j+k})}', \quad (2.2)$$

be an estimator of the spectral density matrix of \mathbf{X}_t , where \odot denotes the Hadamard product. If, for some j and k , $\lambda_{j+k} \notin [-\pi, \pi]$, we take w_n as having period 2π . The function \hat{f}_n can be extended to take values on \mathbb{R} . In order to do so, we need to introduce the following auxiliary function. For a fixed $n > 0$, given $\lambda \in [0, \pi]$, there exists $k_0 \in \mathbb{N}^*$ such that $\lambda_{k_0-1} \leq \lambda \leq \lambda_{k_0}$. Define

$$g(\lambda; n) := \begin{cases} \lambda_{k_0-1}, & \text{if } \lambda - \lambda_{k_0-1} \leq \lambda_{k_0} - \lambda; \\ \lambda_{k_0}, & \text{if } \lambda - \lambda_{k_0-1} > \lambda_{k_0} - \lambda. \end{cases}$$

If $\lambda \in [-\pi, 0)$, we define $g(\lambda; n) := g(-\lambda; n)$. Then, if λ is not a Fourier frequency, we extend (2.2) by setting

$$\hat{f}_n(\lambda) := \hat{f}_n(g(\lambda; n)).$$

Expression (2.2) extended in this way is a multivariate extension of the univariate smoothed periodogram function. In (2.2), we allow the use of different weight functions for different components (including cross spectrum components) of the spectral density matrix. In this manner

we accommodate the necessity, often observed in practice, to model different characteristics present in the spectral density matrix by using different weight functions. We refrain from presenting further discussion on the different types of weight functions in the literature, since they can be found in most text books on the subject. See, for example, Priestley (1981) and references therein.

In practice, at zero frequency, we use a slightly different estimative, namely,

$$\hat{f}_n(0) := \operatorname{Re} \left[W_n(0) \odot w_n(\lambda_1) \overline{w_n(\lambda_1)}' + 2 \sum_{k=1}^{\ell(n)} W_n(k) \odot w_n(\lambda_{k+1}) \overline{w_n(\lambda_{k+1})}' \right].$$

Considering (1.1), under conditions **A1** to **A4**, the smoothed periodogram is $n^{1/2}$ -consistent for the spectral density matrix whenever $\mathbf{d} \in [-1/2, 0]^q$. See, for instance, Grenander and Roseblatt (1953) and Priestley (1981).

3 Smoothed Gaussian Semiparametric Estimation

From the local form of the spectral density at zero frequency (2.1) replaced in the frequency domain Gaussian log-likelihood localized at the origin, Shimotsu (2007) proposed a semiparametric estimator for the fractional differencing parameter \mathbf{d} based on the objective function

$$R(\mathbf{d}) := \log(\det\{\tilde{G}(\mathbf{d})\}) - 2 \sum_{k=1}^q d_k \frac{1}{m} \sum_{j=1}^m \log(\lambda_j), \quad (3.1)$$

where

$$\tilde{G}(\mathbf{d}) := \frac{1}{m} \sum_{j=1}^m \operatorname{Re}[\Lambda_j(\mathbf{d})^{-1} I_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}']. \quad (3.2)$$

with $\Lambda_j(\mathbf{d})$ defined in (2.1). The estimator of \mathbf{d} is then given by

$$\tilde{\mathbf{d}} = \arg \min_{\mathbf{d} \in \Theta} \{R(\mathbf{d})\}, \quad (3.3)$$

where the space of admissible estimates is of the form $\Theta = [-1/2 + \epsilon_1, 1/2 - \epsilon_2]$, for arbitrarily small $\epsilon_i > 0$, $i = 1, 2$, henceforth fixed. In Section 6 we shall denote the estimator (3.3) by Sh . Shimotsu (2007) shows that the estimator based on the objective function (3.1) and (3.2) is consistent under mild conditions. The proof, however, is complicated and involves somewhat delicate results on the periodogram function I_n . A natural question is what happens if the periodogram I_n is replaced by a consistent estimator of the spectral density function? As we shall see later, replacing the periodogram I_n by a consistent estimator of the spectral density is sufficient to guarantee the consistency of the estimator (3.3) in the case of long-range dependence, $\mathbf{d} \in (0, 1/2)^q$. For the case $\mathbf{d} \in (-1/2, 0)^q$, surprisingly, consistency alone is not enough, but n^β -consistency, for some $\beta \in (0, 1)$, guarantees the consistency of the estimator for $\mathbf{d} \in (-\beta/2, 1/2)^q$. Under extra regularity conditions on f_n , the n^β -consistency can be relaxed (see Theorem 4.2). Proofs are significantly simpler than the original ones.

Let f_n be an arbitrary estimator for the spectral density function f . We are interested in estimators based on objective functions of the form

$$S(\mathbf{d}) := \log(\det\{\hat{G}(\mathbf{d})\}) - 2 \sum_{k=1}^q d_k \frac{1}{m} \sum_{j=1}^m \log(\lambda_j), \quad (3.4)$$

with

$$\hat{G}(\mathbf{d}) := \frac{1}{m} \sum_{j=1}^m \operatorname{Re}[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}']. \quad (3.5)$$

Notice that (3.4) is just (3.1) with the periodogram I_n in (3.2) replaced by f_n . The estimator of \mathbf{d} is then defined analogously as

$$\hat{\mathbf{d}} = \arg \min_{\mathbf{d} \in \Theta} \{S(\mathbf{d})\}. \quad (3.6)$$

The estimator (3.6) based on the smoothed periodogram shall be denoted in Section 6 by SSh .

Before proceeding with the results, we shall establish some notation. Let $\{\mathbf{X}_t\}_{t=0}^{\infty}$ be a q -dimensional process specified by (1.1) and let $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ such that $\mathbf{X}_t - \mathbb{E}(\mathbf{X}_t) = \sum_{k=0}^{\infty} A_k \varepsilon_{t-k}$. We define a function A by setting

$$A(\lambda) := \sum_{k=0}^{\infty} A_k e^{ik\lambda}. \quad (3.7)$$

The periodogram function associated to $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ is denoted by I_{ε} , that is,

$$I_{\varepsilon}(\lambda) := w_{\varepsilon}(\lambda) \overline{w_{\varepsilon}(\lambda)}', \quad \text{where} \quad w_{\varepsilon}(\lambda) := \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n \varepsilon_t e^{it\lambda}. \quad (3.8)$$

For a matrix M , we shall denote the r -th row and the s -th column of M by $(M)_{r.}$ and $(M)_{.s}$, respectively.

4 Consistency of the estimator

Let $\{\mathbf{X}_t\}_{t=0}^{\infty}$ be a q -dimensional process specified by (1.1) and f be its spectral density matrix. Suppose that the spectral density matrix of the weakly stationary process $\{\mathbf{Y}_t\}_{t=0}^{\infty}$ in (1.1) satisfy $f_{\mathbf{Y}}(\lambda) \sim G_0$ for a real, symmetric and positive definite matrix $G_0 = (G_0^{rs})_{r,s=1}^q$. Let $\mathbf{d}_0 = (d_1^0, \dots, d_q^0)'$ be the true fractional differencing vector parameter and assume that the following assumptions are satisfied:

B1. As $\lambda \rightarrow 0^+$,

$$f_{rs}(\lambda) = e^{i\pi(d_r^0 - d_s^0)/2} G_0^{rs} \lambda^{-d_r^0 - d_s^0} + o(\lambda^{-d_r^0 - d_s^0}), \quad \text{for all } r, s = 1, \dots, q.$$

B2. Denoting the sup-norm by $\|\cdot\|_{\infty}$, assume that

$$\mathbf{X}_t - \mathbb{E}(\mathbf{X}_t) = \sum_{k=0}^{\infty} A_k \varepsilon_{t-k}, \quad \sum_{k=0}^{\infty} \|A_k\|_{\infty}^2 < \infty, \quad (4.1)$$

where $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ is a process such that

$$\mathbb{E}(\varepsilon_t | \mathcal{F}_{t-1}) = 0 \quad \text{and} \quad \mathbb{E}(\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}) = \mathbf{I}_q, \quad \text{a.s.}$$

for all $t \in \mathbb{Z}$, where \mathbf{I}_q is the $q \times q$ identity matrix and \mathcal{F}_t denotes the σ -field generated by $\{\varepsilon_s, s \leq t\}$. Also assume that there exist a scalar random variable ξ and a constant $K > 0$ such that $\mathbb{E}(\xi^2) < \infty$ and $\mathbb{P}(\|\varepsilon_t\|_{\infty}^2 > \eta) \leq K \mathbb{P}(\xi^2 > \eta)$, for all $\eta > 0$.

B3. In a neighborhood $(0, \delta)$ of the origin, A given by (3.7) is differentiable and, as $\lambda \rightarrow 0^+$,

$$\frac{\partial}{\partial \lambda} (\overline{A(\lambda)})'_{r.} = O(\lambda^{-1} \|(\overline{A(\lambda)})'_{r.}\|_{\infty}).$$

B4. As $n \rightarrow \infty$,

$$\frac{1}{m} + \frac{m}{n} \rightarrow 0.$$

Remark 4.1. Assumptions **B1-B4** are the same as in Shimotsu (2007) and are multivariate extensions of the assumptions made in Robinson (1995b) and analogous to the ones used in Robinson (1995a) and Lobato (1999). Assumption **B1** describes the true spectral density matrix behavior at the origin. Notice that, since $\lim_{\lambda \rightarrow 0^+} e^{i\lambda} - 1 = 0$, replacing $e^{i\pi(d_r^0 - d_s^0)/2}$ by $e^{i(\pi - \lambda)(d_r^0 - d_s^0)/2}$ makes no difference. Assumption **B2** regards the causal representation of \mathbf{X}_t , and more specifically, the behavior of the innovation process which is assumed to be a not necessarily uncorrelated square integrable martingale difference uniformly dominated (in probability) by a scalar random variable with finite second moment. Assumption **B3** is a regularity condition (also imposed in Fox and Taqqu, 1986 and Giraitis and Surgailis, 1990, among others, in the parametric case) and will be useful in proving Lemmas 4.1 and 4.2 below. Assumption **B4** is minimal but necessary since m must go to infinity for consistency, but slower than n in view of Assumption **B1**, which specifies f only at a neighborhood of the zero frequency (recall that $\lambda_j = 2\pi j/n$, $j = 1, \dots, m$).

Observe that assumptions **B1-B4** are only concerned to the behavior of the spectral density matrix on a neighborhood of the origin and, apart from integrability (implied by the process' weakly stationarity property), no assumption whatsoever is made on the spectral density matrix behavior outside this neighborhood. For $\beta \in [0, 1]$ and $q \in \mathbb{N}^*$, let

$$\Omega_\beta := \left[-\frac{\beta}{2}, \frac{1}{2} \right]^q \cap \left[-\frac{1}{2} + \epsilon_1, \frac{1}{2} - \epsilon_2 \right]^q, \quad (4.2)$$

for $\epsilon_i > 0, i \in \{1, 2\}$.

Lemma 4.1 establishes the consistency of $\widehat{G}(\mathbf{d}_0)$ given in (3.5) under the assumption of n^β -consistency of f_n and will be useful in proving Theorem 4.1. The proofs of all results in the paper, due to their lengths, are postponed to the Appendix A.

Lemma 4.1. *Let $\{\mathbf{X}_t\}_{t=0}^\infty$ be a q -dimensional process specified by (1.1) and f be its spectral density matrix. Let f_n be a n^β -consistent estimator for f . If $\mathbf{d}_0 \in \Omega_\beta$, then*

$$\widehat{G}(\mathbf{d}_0) = G_0 + o_{\mathbb{P}}(1).$$

Theorem 4.1 establishes the consistency of $\widehat{\mathbf{d}}$, given in (3.6), under assumptions **B1-B4** and assuming n^β -consistency of the spectral density function estimator.

Theorem 4.1. *Let $\{\mathbf{X}_t\}_{t=0}^\infty$ be a q -dimensional process specified by (1.1) and f be its spectral density matrix. Let f_n be a n^β -consistent estimator of f , for $\beta \in [0, 1]$, and let $\widehat{\mathbf{d}}$ be as in (3.6). Assume that assumptions **B1-B4** hold and let $\mathbf{d}_0 \in \Omega_\beta$. Then, $\widehat{\mathbf{d}} \xrightarrow{\mathbb{P}} \mathbf{d}_0$, as $n \rightarrow \infty$.*

Lemma 4.2 will be useful in proving Theorem 4.2.

Lemma 4.2. *Let $\{\mathbf{X}_t\}_{t=0}^\infty$ be a q -dimensional process specified by (1.1) and f be its spectral density matrix. Let f_n be an estimator of f , and consider the estimator $\widehat{\mathbf{d}}$ based on f as in (3.6). Assume that assumptions **B1-B4** hold. Suppose that f_n satisfies*

$$\mathbb{E} \left(n^{-d_r^0 - d_s^0} \left| f_n^{rs}(\lambda_j) - (A(\lambda_j))_{r.} I_\epsilon(\lambda_j) (\overline{A(\lambda_j)})'_{.s} \right| \right) = O(j^{-d_r^0 - d_s^0 - \gamma}), \quad \text{as } n \rightarrow \infty, \quad (4.3)$$

for some $\gamma > 0$, for all $r, s \in \{1, \dots, q\}$ and $\mathbf{d}_0 \in \Theta$, where A and I_ϵ are given by (3.7) and (3.8), respectively. Then, for $1 \leq u < v \leq m$,

$$\max_{r, s \in \{1, \dots, q\}} \left\{ \sum_{j=u}^v e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} f_n^{rs}(\lambda_j) - G_0^{rs} \right\} = \mathcal{A}_{uv} + \mathcal{B}_{uv},$$

where \mathcal{A}_{uv} and \mathcal{B}_{uv} satisfy

$$\mathbb{E}(|\mathcal{A}_{uv}|) = O(v^{1-\gamma}) \quad \text{and} \quad \max_{1 \leq u < v \leq m} \{ |v^{-1} \mathcal{B}_{uv}| \} = o_{\mathbb{P}}(1).$$

Theorem 4.2, we derive a necessary condition for the consistency of $\hat{\mathbf{d}}$ given in (3.6), when the consistency condition on f_n is relaxed.

Theorem 4.2. *Let $\{\mathbf{X}_t\}_{t=0}^\infty$ be a q -dimensional process specified by (1.1) and f be its spectral density matrix. Let f_n be an estimator of f , and consider the estimator $\hat{\mathbf{d}}$, based on f_n , given in (3.6). Assume that assumptions **B1-B4** hold. Suppose that f_n satisfies (4.3), for some $\gamma > 0$, for all $r, s \in \{1, \dots, q\}$ and $\mathbf{d}_0 \in \Theta$, where A and I_ε are given by (3.7) and (3.8), respectively. Then, $\hat{\mathbf{d}} \xrightarrow{\mathbb{P}} \mathbf{d}_0$, as $n \rightarrow \infty$.*

Corollary 4.1 establishes the consistency of the estimator (3.6) based on the smoothed periodogram under assumptions **A1-A4**. We shall also need an extra minimal condition involving the rate of convergence between m in (3.4) and $\ell(n)$.

A5. As $n \rightarrow \infty$, $\ell(n)/m \rightarrow 0$.

Corollary 4.1. *Let $\{\mathbf{X}_t\}_{t=0}^\infty$ be a q -dimensional process specified by (1.1) and f be its spectral density matrix. Assume that assumptions **B1-B4** hold. Let $\{W_n^{ij}(k)\}_{i,j=1}^q$ be a sequence of weights satisfying assumptions **A1-A5** and let f_n be the respective smoothed periodogram given in (2.2). For $\mathbf{d}_0 \in \Theta$, consider the estimator (3.6). Then, $\hat{\mathbf{d}} \xrightarrow{\mathbb{P}} \mathbf{d}_0$, as $n \rightarrow \infty$.*

We observe that assumption **A5** is not necessary when $\mathbf{d}_0 \in (0, 1/2)^q$, as it is clear from the proof of Corollary 4.1.

Remark 4.2. Recall that the smoothed periodogram is $n^{1/2}$ -consistent (under **A1-A4**) for $\mathbf{d}_0 \in [-1/2, 0]^q$, so that Theorem 4.1 applies and we conclude that the estimator (3.6) is consistent for $\mathbf{d}_0 \in \Omega_{1/2} \cap (-1/2, 0] \subset [-1/4, 0]$. In contrast, as shown in Corollary 4.1, Theorem 4.2 also holds and establishes the consistency of (3.6) for all admissible parameter in Θ , under **A1-A5**. The latter is a far more interesting result, but to obtain it, we had to show that condition (4.3) holds (see the proof of Corollary 4.1), which required some cumbersome computations. The former result, in contrast, was automatically obtained from the properties of the spectral density estimator. Furthermore, Theorem 4.2 does not assume consistency of the underline spectral density estimator. In fact, the periodogram itself satisfies condition (4.3) (see lemma 1(a) in Shimotsu, 2007).

5 Asymptotic normality of the estimator

In this section we present a sufficient condition for the asymptotic normality of the GSE given by (3.6), under similar assumptions to Shimotsu (2007), with f_n an estimator of the spectral density function satisfying a single regularity condition. The asymptotic distribution of the estimator (3.6) will be the same as (3.3), established by Shimotsu (2007).

Again, let $\{\mathbf{X}_t\}_{t=0}^\infty$ be a q -dimensional process specified by (1.1) and f be its spectral density matrix. Suppose that the spectral density matrix of the weakly stationary process $\{\mathbf{Y}_t\}_{t=0}^\infty$ in (1.1) satisfies $f_{\mathbf{Y}}(\lambda) \sim G_0$ for a real, symmetric and positive definite matrix $G_0 = (G_0^{rs})_{r,s=1}^q$. Let $\mathbf{d}_0 = (d_1^0, \dots, d_q^0)'$ be the true fractional differencing vector parameter. Assume that the following assumptions are satisfied

C1. For $\alpha \in (0, 2]$ and $r, s \in \{1, \dots, q\}$,

$$f_{rs}(\lambda) = e^{i(\pi-\lambda)(d_r^0-d_s^0)/2} \lambda^{-d_r^0-d_s^0} G_0^{rs} + O(\lambda^{-d_r^0-d_s^0+\alpha}), \quad \text{as } \lambda \rightarrow 0^+.$$

C2. Assumption **B2** holds and the process $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ has finite fourth moment.

C3. Assumption **B3** holds.

C4. For any $\delta > 0$,

$$\frac{1}{m} + \frac{m^{1+2\alpha} \log(m)^2}{n^{2\alpha}} + \frac{\log(n)}{m^\delta} \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

C5. There exists a finite real matrix M such that

$$\Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) = M + o(1), \quad \text{as } \lambda_j \rightarrow 0.$$

Remark 5.1. Assumption **C1** is a smoothness condition often imposed in spectral analysis. Compared to assumption 1 in Robinson (1995a), assumption **C1** is slightly more restrictive. It is satisfied by certain VARFIMA processes. Assumption **C2** imposes that the process $\{\mathbf{X}_t\}_{t \in \mathbb{N}^*}$ is linear with finite fourth moment. This restriction in the innovation process is necessary in the proof of Theorem 5.1 in applying a CLT to a certain martingale difference derived from a quadratic form involving $\{\varepsilon_t\}_{t \in \mathbb{Z}}$, which must have finite variance. Assumption **C4** is the same as assumption 4' in Shimotsu (2007) and is slightly stronger than the ones imposed in Robinson (1995b) and Lobato (1999) (see Shimotsu, 2007 p.283 for a discussion). It implies that $(m/n)^b = o(m^{-\frac{b}{2\alpha}} \log(m)^{-\frac{b}{\alpha}})$, for $b \neq 0$. Assumption **C5** is the same as assumption 5' in Shimotsu (2007) and is a mild regularity condition in the degree of approximation of $A(\lambda_j)$ by $\Lambda_j(\mathbf{d}_0)$. It is satisfied by general VARFIMA processes.

The next lemma will be useful in proving Theorem 5.1. The proofs of the results in this section are presented in Appendix A.

Lemma 5.1. Let $\{\mathbf{X}_t\}_{t=0}^\infty$ be a q -dimensional process specified by (1.1) and f be its spectral density matrix. Let f_n be an estimator of f , and consider the estimator $\hat{\mathbf{d}}$, based on f_n , given in (3.6). Assume that assumptions **C1-C5** hold. Suppose that f_n satisfies

$$\sum_{j=1}^v \left| f_n^{rs}(\lambda_j) - (A(\lambda_j))_{r \cdot} I_\varepsilon(\lambda_j) (\overline{A(\lambda_j)})'_{\cdot s} \right| = o_P \left(\frac{n^{d_r^0+d_s^0}}{\log(v) v^{d_r^0+d_s^0-1/2}} \right), \quad (5.1)$$

for all $r, s \in \{1, \dots, q\}$, $1 \leq v \leq m$ and $\mathbf{d}_0 \in \Theta$, where I_ε and A are defined in (3.8) and (3.7), respectively. Then,

(a) uniformly in $1 \leq v \leq m$,

$$\max_{r, s \in \{1, \dots, q\}} \left\{ \sum_{j=1}^v e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} \left[f_n^{rs}(\lambda_j) - (A(\lambda_j))_{r \cdot} I_\varepsilon(\lambda_j) (\overline{A(\lambda_j)})'_{\cdot s} \right] \right\} = o_P \left(\frac{v^{1/2}}{\log(v)} \right); \quad (5.2)$$

(b) uniformly in $1 \leq v \leq m$,

$$\max_{r, s \in \{1, \dots, q\}} \left\{ \sum_{j=1}^v e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} f_n^{rs}(\lambda_j) - G_0^{rs} \right\} = O_P \left(\frac{v^{\alpha+1}}{n^\alpha} + v^{1/2} \log(v) \right). \quad (5.3)$$

The next theorem presents a necessary condition for the asymptotic normality of the GSE given in (3.6). We notice that the variance-covariance matrix of the limiting distribution is the same as the estimator in (3.3), as derived in Shimotsu (2007).

Theorem 5.1. *Let $\{\mathbf{X}_t\}_{t=0}^{\infty}$ be a q -dimensional process specified by (1.1) and f be its spectral density matrix. Let f_n be an estimator of f , and consider the estimator $\hat{\mathbf{d}}$, based on f_n , given in (3.6). Assume that assumptions **C1-C5** hold. Suppose that f_n satisfies (5.1), for all $r, s \in \{1, \dots, q\}$ and $\mathbf{d}_0 \in \Theta$. If $\hat{\mathbf{d}} \xrightarrow{\mathbb{P}} \mathbf{d}_0$, for $\mathbf{d}_0 \in \Theta$, then*

$$m^{1/2}(\hat{\mathbf{d}} - \mathbf{d}_0) \xrightarrow{d} N(\mathbf{0}, \Omega^{-1}),$$

as n tends to infinity, where

$$\Omega := 2 \left[G_0 \odot G_0^{-1} + \mathbf{I}_q + \frac{\pi^2}{4} (G_0 \odot G_0^{-1} - \mathbf{I}_q) \right],$$

with \mathbf{I}_q the $q \times q$ identity matrix.

6 Monte Carlo Simulation Study

In this section we present a Monte Carlo simulation study to assess the finite sample performance of the proposed estimator given in (3.6) (denoted *SSh*). For comparison purposes, we also calculate the estimator (3.3) (denoted *Sh*). All Monte Carlo simulations are based on time series of fixed sample size $n = 1,000$ obtained from bidimensional Gaussian VARFIMA(0, \mathbf{d} , 0) processes for several different parameters \mathbf{d} and correlations $\rho \in \{0, 0.3, 0.6, 0.8\}$. We perform 1,000 replications of each experiment. To generate the time series, we apply the traditional method of truncating the multidimensional infinite moving average representation of the process. The truncation point is fixed in 50,000 for all cases.

In all simulations, we apply the smoothed periodogram (2.2) with the same weights for all spectral density components, given by the so-called Bartlett's window, that is, we use

$$W_n^{ij}(k) := \frac{\sin^2(nk/2)}{2\pi n \sin^2(k/2)}, \quad \text{for all } i, j = 1, 2.$$

The truncation point of the smoothed periodogram function is of the form $\ell(n, \beta) := \lfloor n^\beta \rfloor$, for $\beta \in \{0.7, 0.9\}$, while the truncation point of the estimators *SSh* and *Sh* are both of the form $m := m(n) = \lfloor n^\alpha \rfloor$, for $\alpha \in \{0.65, 0.85\}$.

The routines are implemented in FORTRAN 95 language optimized by using OpenMP directives for parallel computing. All simulations were performed by using the computational resources from the (Brazilian) National Center of Super Computing (CESUP-UFRGS).

Table 6.1 reports the simulation results for $\mathbf{d} \in \{(0.1, 0.4), (0.4, 0.4), (0.2, 0.2), (0.2, 0.3)\}$. Presented are the estimated values (mean), their standard deviations (st.d.) and the mean square error of the estimates (mse).

We observe that in all cases both estimators perform well, but the *SSh* usually present a slightly better performance in terms of mse compared to the *Sh* estimator. In terms of mse, the best combination of (α, β) for the *SSh* estimator is (0.85, 0.9) in all cases. For the *Sh* estimator, we found that $\alpha = 0.85$ produces better results than $\alpha = 0.65$, applied in the simulations presented in Shimotsu (2007). As the correlation in the innovation process increases, the estimated values slightly degrade.

In Table 6.1 we notice that the trade off between variance and bias, an usual feature in smoothing the periodogram, has little effect in the estimated values. In most cases, we observe that the higher the value of β , the higher the variance but smaller the bias of the estimate.

Table 6.1: Simulation results of the estimator (3.6) based on the smoothed periodogram with the Bartlett's spectral window (weights) and estimator (3.3) in VARFIMA(0, \mathbf{d} , 0) processes. Presented are the estimated values (mean), its standard deviation (st.d) and the mean square error of the estimates (mse).

| ρ | Method | β | \hat{d}_i | $\mathbf{d} = (0.1, 0.4)$ | | | | | | $\mathbf{d} = (0.4, 0.4)$ | | | | | |
|--------|--------|---------|-------------|---------------------------|--------|--------|-----------------|--------|--------|---------------------------|---------|--------|-----------------|--------|--------|
| | | | | $\alpha = 0.65$ | | | $\alpha = 0.85$ | | | $\alpha = 0.65$ | | | $\alpha = 0.85$ | | |
| | | | | mean | st.d. | mse | mean | st.d. | mse | mean | st.d. | mse | mean | st.d. | mse |
| 0 | SSH | 0.7 | \hat{d}_1 | 0.1030 | 0.0544 | 0.0030 | 0.0955 | 0.0268 | 0.0007 | 0.4506 | 0.0709 | 0.0076 | 0.4119 | 0.0366 | 0.0015 |
| | | | \hat{d}_2 | 0.4467 | 0.0789 | 0.0084 | 0.4140 | 0.0402 | 0.0018 | 0.4453 | 0.0781 | 0.0082 | 0.4138 | 0.0401 | 0.0018 |
| | | 0.9 | \hat{d}_1 | 0.1047 | 0.0447 | 0.0032 | 0.0953 | 0.0268 | 0.0007 | 0.4312 | 0.0601 | 0.0046 | 0.3944 | 0.0291 | 0.0009 |
| | | | \hat{d}_2 | 0.4230 | 0.0539 | 0.0044 | 0.3949 | 0.0306 | 0.0010 | 0.4228 | 0.0624 | 0.0044 | 0.3948 | 0.0306 | 0.0010 |
| | Sh | - | \hat{d}_1 | 0.1053 | 0.0581 | 0.0020 | 0.0955 | 0.0271 | 0.0005 | 0.4136 | 0.0590 | 0.0020 | 0.3851 | 0.0277 | 0.0008 |
| | | | \hat{d}_2 | 0.3955 | 0.0606 | 0.0022 | 0.3822 | 0.0292 | 0.0006 | 0.3955 | 0.0606 | 0.0021 | 0.3822 | 0.0292 | 0.0007 |
| 0.3 | SSH | 0.7 | \hat{d}_1 | 0.1100 | 0.0513 | 0.0027 | 0.0982 | 0.0257 | 0.0007 | 0.4497 | 0.0648 | 0.0067 | 0.4118 | 0.0346 | 0.0013 |
| | | | \hat{d}_2 | 0.4468 | 0.0741 | 0.0077 | 0.4148 | 0.0384 | 0.0017 | 0.4478 | 0.0714 | 0.0074 | 0.4142 | 0.0375 | 0.0016 |
| | | 0.9 | \hat{d}_1 | 0.1076 | 0.0525 | 0.0028 | 0.0960 | 0.0256 | 0.0007 | 0.4299 | 0.0555 | 0.0040 | 0.3942 | 0.0277 | 0.0008 |
| | | | \hat{d}_2 | 0.4243 | 0.0587 | 0.0040 | 0.3958 | 0.0288 | 0.0008 | 0.4253 | 0.0581 | 0.0040 | 0.3952 | 0.0286 | 0.0008 |
| | Sh | - | \hat{d}_1 | 0.1038 | 0.0542 | 0.0030 | 0.0947 | 0.0259 | 0.0007 | 0.4109 | 0.0550 | 0.0031 | 0.3846 | 0.0265 | 0.0009 |
| | | | \hat{d}_2 | 0.3985 | 0.0572 | 0.0033 | 0.3834 | 0.0273 | 0.0010 | 0.3991 | 0.0570 | 0.0032 | 0.3828 | 0.0273 | 0.0010 |
| 0.6 | SSH | 0.7 | \hat{d}_1 | 0.1298 | 0.0483 | 0.0032 | 0.1073 | 0.0248 | 0.0007 | 0.4499 | 0.0566 | 0.0057 | 0.4125 | 0.0310 | 0.0011 |
| | | | \hat{d}_2 | 0.4500 | 0.0665 | 0.0069 | 0.4182 | 0.0361 | 0.0016 | 0.4503 | 0.0607 | 0.0062 | 0.4143 | 0.0327 | 0.0013 |
| | | 0.9 | \hat{d}_1 | 0.1167 | 0.0467 | 0.0025 | 0.0990 | 0.0234 | 0.0006 | 0.4290 | 0.0485 | 0.0032 | 0.3943 | 0.0250 | 0.0007 |
| | | | \hat{d}_2 | 0.4274 | 0.0525 | 0.0035 | 0.3981 | 0.0262 | 0.0007 | 0.4280 | 0.0505 | 0.0033 | 0.3953 | 0.0252 | 0.0007 |
| | Sh | - | \hat{d}_1 | 0.1034 | 0.0478 | 0.0023 | 0.0939 | 0.0239 | 0.0006 | 0.4077 | 0.04077 | 0.0024 | 0.3840 | 0.0239 | 0.0008 |
| | | | \hat{d}_2 | 0.4029 | 0.0501 | 0.0025 | 0.3858 | 0.0241 | 0.0008 | 0.4037 | 0.0498 | 0.0025 | 0.3836 | 0.0240 | 0.0008 |
| 0.8 | SSH | 0.7 | \hat{d}_1 | 0.1570 | 0.0512 | 0.0059 | 0.1233 | 0.0278 | 0.0013 | 0.4506 | 0.0535 | 0.0054 | 0.4130 | 0.0290 | 0.0010 |
| | | | \hat{d}_2 | 0.4669 | 0.0647 | 0.0087 | 0.4295 | 0.0374 | 0.0023 | 0.4512 | 0.0557 | 0.0057 | 0.4142 | 0.0298 | 0.0011 |
| | | 0.9 | \hat{d}_1 | 0.1307 | 0.0458 | 0.0030 | 0.1055 | 0.0231 | 0.0006 | 0.4288 | 0.0451 | 0.0029 | 0.3946 | 0.0231 | 0.0006 |
| | | | \hat{d}_2 | 0.4372 | 0.0509 | 0.0040 | 0.4037 | 0.0258 | 0.0007 | 0.4290 | 0.0464 | 0.0030 | 0.3953 | 0.0230 | 0.0006 |
| | Sh | - | \hat{d}_1 | 0.1060 | 0.0448 | 0.0020 | 0.0946 | 0.0218 | 0.0005 | 0.4066 | 0.0448 | 0.0020 | 0.3838 | 0.0221 | 0.0008 |
| | | | \hat{d}_2 | 0.4071 | 0.0465 | 0.0022 | 0.3882 | 0.0221 | 0.0006 | 0.4056 | 0.0458 | 0.0021 | 0.3840 | 0.0218 | 0.0007 |
| ρ | Method | β | \hat{d}_i | $\mathbf{d} = (0.2, 0.2)$ | | | | | | $\mathbf{d} = (0.2, 0.3)$ | | | | | |
| | | | | $\alpha = 0.65$ | | | $\alpha = 0.85$ | | | $\alpha = 0.65$ | | | $\alpha = 0.85$ | | |
| | | | | mean | st.d. | mse | mean | st.d. | mse | mean | st.d. | mse | mean | st.d. | mse |
| 0 | SSH | 0.7 | \hat{d}_1 | 0.2085 | 0.0570 | 0.0033 | 0.1946 | 0.0277 | 0.0008 | 0.2085 | 0.0569 | 0.0033 | 0.1946 | 0.0277 | 0.0008 |
| | | | \hat{d}_2 | 0.1994 | 0.0572 | 0.0033 | 0.1950 | 0.0294 | 0.0009 | 0.3164 | 0.0643 | 0.0044 | 0.3000 | 0.0321 | 0.0010 |
| | | 0.9 | \hat{d}_1 | 0.2087 | 0.0565 | 0.0033 | 0.1924 | 0.0269 | 0.0008 | 0.2087 | 0.0565 | 0.0033 | 0.1924 | 0.0269 | 0.0008 |
| | | | \hat{d}_2 | 0.1989 | 0.0567 | 0.0032 | 0.1926 | 0.0286 | 0.0009 | 0.3078 | 0.0580 | 0.0034 | 0.2919 | 0.0289 | 0.0009 |
| | Sh | - | \hat{d}_1 | 0.2072 | 0.0581 | 0.0034 | 0.1915 | 0.0271 | 0.0008 | 0.2072 | 0.0581 | 0.0034 | 0.1915 | 0.0271 | 0.0008 |
| | | | \hat{d}_2 | 0.1925 | 0.0601 | 0.0037 | 0.1902 | 0.0291 | 0.0009 | 0.2937 | 0.0603 | 0.0037 | 0.2860 | 0.0291 | 0.0010 |
| 0.3 | SSH | 0.7 | \hat{d}_1 | 0.2072 | 0.0530 | 0.0029 | 0.1945 | 0.0264 | 0.0007 | 0.2097 | 0.0530 | 0.0029 | 0.1951 | 0.0264 | 0.0007 |
| | | | \hat{d}_2 | 0.2014 | 0.0539 | 0.0029 | 0.1952 | 0.0275 | 0.0008 | 0.3171 | 0.0601 | 0.0039 | 0.3000 | 0.0301 | 0.0009 |
| | | 0.9 | \hat{d}_1 | 0.2073 | 0.0527 | 0.0028 | 0.1923 | 0.0257 | 0.0007 | 0.2085 | 0.0527 | 0.0028 | 0.1924 | 0.0257 | 0.0007 |
| | | | \hat{d}_2 | 0.2009 | 0.0538 | 0.0029 | 0.1927 | 0.0268 | 0.0008 | 0.3093 | 0.0547 | 0.0031 | 0.2922 | 0.0270 | 0.0008 |
| | Sh | - | \hat{d}_1 | 0.2050 | 0.0543 | 0.0030 | 0.1912 | 0.0260 | 0.0008 | 0.2051 | 0.0542 | 0.0030 | 0.1910 | 0.0259 | 0.0008 |
| | | | \hat{d}_2 | 0.1955 | 0.0567 | 0.0032 | 0.1907 | 0.0272 | 0.0008 | 0.2967 | 0.0568 | 0.0032 | 0.2867 | 0.0272 | 0.0009 |
| 0.6 | SSH | 0.7 | \hat{d}_1 | 0.2058 | 0.0464 | 0.0022 | 0.1946 | 0.0239 | 0.0006 | 0.2138 | 0.0470 | 0.0024 | 0.1970 | 0.0240 | 0.0006 |
| | | | \hat{d}_2 | 0.2039 | 0.0473 | 0.0022 | 0.1952 | 0.0242 | 0.0006 | 0.3178 | 0.0525 | 0.0031 | 0.3001 | 0.0267 | 0.0007 |
| | | 0.9 | \hat{d}_1 | 0.2057 | 0.0462 | 0.0022 | 0.1923 | 0.0233 | 0.0006 | 0.2095 | 0.0461 | 0.0022 | 0.1929 | 0.0233 | 0.0006 |
| | | | \hat{d}_2 | 0.2035 | 0.0473 | 0.0023 | 0.1928 | 0.0236 | 0.0006 | 0.3111 | 0.0481 | 0.0024 | 0.2926 | 0.0239 | 0.0006 |
| | Sh | - | \hat{d}_1 | 0.2025 | 0.0478 | 0.0023 | 0.1909 | 0.0235 | 0.0006 | 0.2030 | 0.0477 | 0.0023 | 0.1904 | 0.0235 | 0.0006 |
| | | | \hat{d}_2 | 0.1993 | 0.0496 | 0.0025 | 0.1911 | 0.0239 | 0.0007 | 0.3005 | 0.0496 | 0.0025 | 0.2877 | 0.0239 | 0.0007 |
| 0.8 | SSH | 0.7 | \hat{d}_1 | 0.2054 | 0.0430 | 0.0019 | 0.1948 | 0.0222 | 0.0005 | 0.2192 | 0.0448 | 0.0024 | 0.1998 | 0.0228 | 0.0005 |
| | | | \hat{d}_2 | 0.2049 | 0.0436 | 0.0019 | 0.1952 | 0.0221 | 0.0005 | 0.3208 | 0.0492 | 0.0028 | 0.3014 | 0.0249 | 0.0006 |
| | | 0.9 | \hat{d}_1 | 0.2051 | 0.0428 | 0.0019 | 0.1925 | 0.0216 | 0.0005 | 0.2117 | 0.0428 | 0.0020 | 0.1939 | 0.0216 | 0.0005 |
| | | | \hat{d}_2 | 0.2046 | 0.0435 | 0.0019 | 0.1928 | 0.0215 | 0.0005 | 0.3131 | 0.0445 | 0.0021 | 0.2933 | 0.0219 | 0.0005 |
| | Sh | - | \hat{d}_1 | 0.2016 | 0.0445 | 0.0020 | 0.1909 | 0.0218 | 0.0006 | 0.2026 | 0.0444 | 0.0020 | 0.1903 | 0.0218 | 0.0006 |
| | | | \hat{d}_2 | 0.2008 | 0.0455 | 0.0021 | 0.1912 | 0.0217 | 0.0005 | 0.3024 | 0.0456 | 0.0021 | 0.2883 | 0.0217 | 0.0006 |

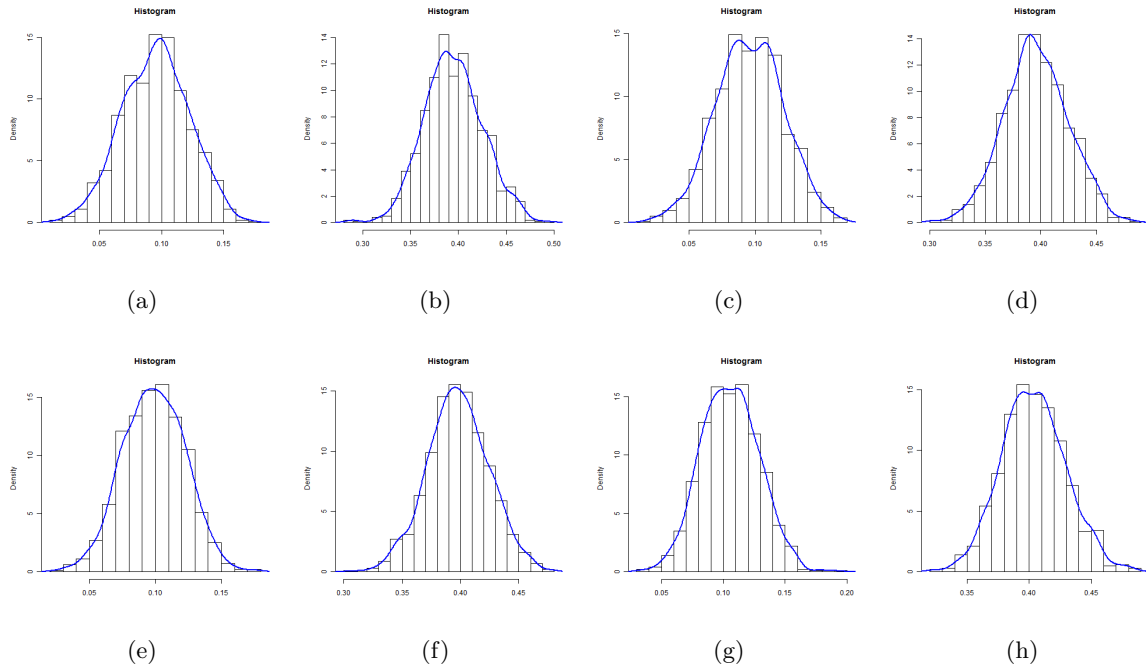


Figure 6.1: Histogram and kernel density of the *SSh* estimated values of $\mathbf{d}_0 = (0.1, 0.4)$ for (a) and (b), $\rho = 0$, (c) and (d), $\rho = 0.3$, (e) and (f), $\rho = 0.6$ and (g) and (h), $\rho = 0.8$.

The variance of the estimator responds stronger to changes in α than in β in the opposite direction, that is, the higher the α , the smaller the variance. The estimated value of \mathbf{d} by the *SSh* seems to respond stronger to changes in α than in β .

Figure 6.1 presents the histogram and kernel density estimator of the *SSh* estimated values for $\mathbf{d}_0 = (0.1, 0.4)$ when $\alpha = 0.85$, $\beta = 0.9$. Figures 6.1(a) and 6.1(b) correspond to $\rho = 0$, Figures 6.1(c) and 6.1(d) to $\rho = 0.3$, Figures 6.1(e) and 6.1(f) to $\rho = 0.6$ and Figures 6.1(g) and 6.1(h) to $\rho = 0.8$. Notice that the resemblance of the histograms to the normal distribution is greater for $\mathbf{d}_0^{(2)} = 0.4$ than to 0.1. At this moment, we were not able to prove the asymptotic normality of the *SSh* estimator with the smoothed periodogram by direct verification of (5.1). However we conjecture that this is the case and Figure 6.1 supports this opinion.

7 Conclusions

In this work we propose and analyze a class of Gaussian semiparametric estimators of multivariate long-range dependent processes. The work is motivated by the semiparametric methodology presented in Shimotsu (2007). More specifically, we propose a class of estimators based on the method studied in Shimotsu (2007) by substituting the periodogram applied there for an arbitrary spectral density estimator. We analyze two frameworks. First we assume that the spectral density estimator is consistent for the spectral density estimator and we show that the proposed semiparametric estimator is also consistent under mild conditions. Second, we relax the consistency condition and derive necessary conditions for the consistency and asymptotic normality of the proposed estimator. We show that the variance-covariance matrix of the limiting distribution is the same as the one derived in Shimotsu (2007), under the same conditions imposed in the process.

In order to assess the finite sample performance and illustrate the usefulness of the estimator, we perform a Monte Carlo simulation based on VARFIMA(0, \mathbf{d} , 0) processes. We applied the so-called smoothed periodogram (which is shown to satisfy the consistency conditions imposed

in the theoretical results) with the Bartlett's weight function as the spectral density estimator. For comparison we also compute the estimator proposed in Shimotsu (2007). Both estimators perform great but the one proposed here present generally better results.

The assumptions required in the asymptotic theory are mild ones and are commonly applied in the literature. The semiparametric methodology present several advantages compared to the parametric framework such as weaker distributional assumptions, robustness with respect to misspecification of the short run dynamics of the process and efficiency. The theory includes the fractionally integrated processes as well as the class of VARFIMA processes.

Acknowledgements

G. Pumi was partially supported by CAPES/Fulbright Grant BEX 2910/06-3 and by CNPq-Brazil. S.R.C. Lopes research was partially supported by CNPq-Brazil, by CAPES-Brazil, by Pronex *Probabilidade e Processos Estocásticos* - E-26/170.008/2008 -APQ1 and also by INCT *em Matemática*. The authors are also grateful to the (Brazilian) National Center of Super Computing (CESUP-UFRGS) for the computational resources.

Appendix A: Proofs

In this section we present the proofs of the results in Sections 4 and 5. We establish lemmas and theorems in the same sequence as they appear in the text.

Proof of Lemma 4.1:

By hypothesis, $f_n(\lambda) = f(\lambda) + o_{\mathbb{P}}(n^{-\beta})$. Recalling the definition of Λ_j given in (2.1), we have

$$\begin{aligned} \widehat{G}(\mathbf{d}_0) &= \frac{1}{m} \sum_{j=1}^m \operatorname{Re}[\Lambda_j(\mathbf{d}_0)^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}'] \\ &= \frac{1}{m} \sum_{j=1}^m \operatorname{Re}[\Lambda_j(\mathbf{d}_0)^{-1} (f(\lambda_j) + o_{\mathbb{P}}(n^{-\beta})) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}'] \\ &= G_0 + \sum_{k=1}^q \left(\left[\frac{1}{m} \sum_{j=1}^m \lambda_j^{2d_k^0} \right] o_{\mathbb{P}}(n^{-\beta}) \right) = G_0 + o_{\mathbb{P}}(1), \end{aligned}$$

since

$$\begin{aligned} \left[\frac{1}{m} \sum_{j=1}^m \lambda_j^{2d_k^0} \right] o_{\mathbb{P}}(n^{-\beta}) &= \frac{1}{2d_k^0 + 1} \left(\frac{2\pi m}{n} \right)^{2d_k^0} \left[\frac{2d_k^0 + 1}{m} \sum_{j=1}^m \left(\frac{j}{m} \right)^{2d_k^0} \right] o_{\mathbb{P}}(n^{-\beta}) \\ &= \frac{1}{2d_k^0 + 1} \left(\frac{2\pi m}{n} \right)^{2d_k^0} [O(m^{\beta-1}) + 1] o_{\mathbb{P}}(n^{-\beta}) = o_{\mathbb{P}}(1), \end{aligned}$$

where the second equality follows from lemma 2 in Robinson (1995b), by taking $\gamma = 2d_k^0 + 1 > 0$. The last equality is justified as follows: recalling that $\beta \in (0, 1)$, if $d_k^0 \geq 0$, $(m/n)^{2d_k^0} \rightarrow 0$, and the result is immediate; if $d_k^0 < 0$, $m^{2d_k^0} \rightarrow 0$ and the result follows from the assumption that $\mathbf{d}_0 \in \Omega_{\beta}$, which implies $2d_k^0 + \beta \geq 0$. This completes the proof. \blacksquare

Proof of Theorem 4.1:

Let $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)' := \mathbf{d} - \mathbf{d}_0$ and $L(\mathbf{d}) := S(\mathbf{d}) - S(\mathbf{d}_0)$. Let $0 < \delta < 1/2$ be fixed and let

$$N_{\delta} := \{\mathbf{d} : \|\mathbf{d} - \mathbf{d}_0\|_{\infty} > \delta\}.$$

Let $0 < \epsilon < 1/4$ and define $\Theta_1 := \{\boldsymbol{\theta} : \boldsymbol{\theta} \in [-1/2 + \epsilon, 1/2]^q\}$ and $\Theta_2 = \Omega_\beta \setminus \Theta_1$ (possibly an empty set), where Ω_β is given by (4.2). Following Robinson (1995b) and Shimotsu (2007), we have

$$\begin{aligned} \mathbb{P}(\|\hat{\mathbf{d}} - \mathbf{d}_0\|_\infty > \delta) &\leq \mathbb{P}\left(\inf_{\bar{N}_\delta \cap \Omega_\beta} \{L(\mathbf{d})\} \leq 0\right) \\ &\leq \mathbb{P}\left(\inf_{\bar{N}_\delta \cap \Theta_1} \{L(\mathbf{d})\} \leq 0\right) + \mathbb{P}\left(\inf_{\Theta_2} \{L(\mathbf{d})\} \leq 0\right) := P_1 + P_2 \end{aligned} \quad (\text{A.1})$$

where, for a given set \mathcal{O} , $\bar{\mathcal{O}}$ denotes the closure of \mathcal{O} . We shall first show that $P_1 \rightarrow 0$, as n tends to infinity. Rewrite $L(\mathbf{d})$ as

$$\begin{aligned} L(\mathbf{d}) &= \log(\det\{\widehat{G}(\mathbf{d})\}) - \log(\det\{\widehat{G}(\mathbf{d}_0)\}) - 2 \sum_{k=1}^q \theta_k \frac{1}{m} \sum_{j=1}^m \log(\lambda_j) \\ &= \log(\det\{\widehat{G}(\mathbf{d})\}) - \log(\det\{\widehat{G}(\mathbf{d}_0)\}) + \log\left(\frac{2\pi m}{n}\right)^{-2 \sum_k \theta_k} - \\ &\quad - 2 \sum_{k=1}^q \theta_k \left(\frac{1}{m} \sum_{j=1}^m \log(j) - \log(m)\right) - \sum_{k=1}^q \log(2\theta_k + 1) \\ &= \log(\mathcal{A}(\mathbf{d})) - \log(\mathcal{B}(\mathbf{d})) - \log(\mathcal{A}(\mathbf{d}_0)) + \log(\mathcal{B}(\mathbf{d}_0)) + \mathcal{R}(\mathbf{d}) \\ &= Q_1(\mathbf{d}) - Q_2(\mathbf{d}) + \mathcal{R}(\mathbf{d}), \end{aligned} \quad (\text{A.2})$$

where

$$\begin{aligned} Q_1(\mathbf{d}) &:= \log(\mathcal{A}(\mathbf{d})) - \log(\mathcal{B}(\mathbf{d})), \quad Q_2(\mathbf{d}) := \log(\mathcal{A}(\mathbf{d}_0)) + \log(\mathcal{B}(\mathbf{d}_0)), \\ \mathcal{A}(\mathbf{d}) &:= \left(\frac{2\pi m}{n}\right)^{-2 \sum_k \theta_k} \det\{\widehat{G}(\mathbf{d})\}, \quad \mathcal{B}(\mathbf{d}) := \det\{G_0\} \prod_{k=1}^q \frac{1}{2\theta_k + 1}, \\ \text{and } \mathcal{R}(\mathbf{d}) &:= 2 \sum_{k=1}^q \theta_k \left(\log(m) - \frac{1}{m} \sum_{j=1}^m \log(j)\right) - \sum_{k=1}^q \log(2\theta_k + 1). \end{aligned}$$

By lemma 2 in Robinson (1995b), $\log(m) - m^{-1} \sum_{j=1}^m \log(j) = 1 + O(m^{-1} \log(m))$, so that

$$\mathcal{R}(\mathbf{d}) = \sum_{k=1}^q \theta_k \left[2\theta_k - \log(2\theta_k + 1)\right] + O\left(\frac{\log(m)}{m}\right).$$

Since $x - \log(x+1)$ has a unique global minimum in $(-1, \infty)$ at $x = 0$ and $x - \log(x+1) \geq x^2/4$, for $|x| \leq 1$, it follows that

$$\inf_{\bar{N}_\delta \cap \Theta_1} \{\mathcal{R}(\mathbf{d})\} \geq \frac{1}{4} \left(2 \max_k \{\theta_k\}\right)^2 \geq \delta^2 > 0.$$

As for $Q_1(\mathbf{d})$ and $Q_2(\mathbf{d})$ in (A.2), it suffices to show the existence of a function $h(\mathbf{d}) > 0$ satisfying

$$(i) \sup_{\Theta_1} \{|\mathcal{A}(\mathbf{d}) - h(\mathbf{d})|\} = o_{\mathbb{P}}(1); \quad (ii) h(\mathbf{d}) \geq \mathcal{B}(\mathbf{d}); \quad (iii) h(\mathbf{d}_0) = \mathcal{B}(\mathbf{d}_0),$$

as n goes to infinity, because (ii) implies $\inf_{\Theta_1} \{h(\mathbf{d})\} \geq \inf_{\Theta_1} \{\mathcal{B}(\mathbf{d})\} > 0$, so that, uniformly in Θ_1 ,

$$Q_1(\mathbf{d}) \geq \log(\mathcal{A}(\mathbf{d})) - \log(h(\mathbf{d})) = \log(h(\mathbf{d}) + o_{\mathbb{P}}(1)) - \log(h(\mathbf{d})) = o_{\mathbb{P}}(1),$$

and (iii) implies $Q_2(\mathbf{d}) = \log(h(\mathbf{d}_0) + o_{\mathbb{P}}(1)) - \log(h(\mathbf{d}_0)) = o_{\mathbb{P}}(1)$, from which $P_1 \rightarrow 0$ follows. To show (i), recall that

$$\begin{aligned} \Lambda_j(\mathbf{d})^{-1} &= \text{diag}_{k \in \{1, \dots, q\}} \{\lambda_j^{d_k} e^{i(\lambda_j - \pi)d_k/2}\} = \text{diag}_{k \in \{1, \dots, q\}} \{\lambda_j^{(d_k - d_k^0)} e^{i(\lambda_j - \pi)(d_k - d_k^0)/2} \times \lambda_j^{d_k^0} e^{i(\lambda_j - \pi)d_k^0/2}\} \\ &= \Lambda_j(\mathbf{d} - \mathbf{d}_0)^{-1} \Lambda_j(\mathbf{d}_0)^{-1} = \Lambda_j(\boldsymbol{\theta})^{-1} \Lambda_j(\mathbf{d}_0)^{-1}, \end{aligned}$$

so that we can write

$$\widehat{G}(\mathbf{d}) = \frac{1}{m} \sum_{j=1}^m \operatorname{Re} [\Lambda_j(\boldsymbol{\theta})^{-1} \widehat{G}(\mathbf{d}_0) \overline{\Lambda_j(\boldsymbol{\theta})^{-1}}'] \quad (\text{A.3})$$

and

$$\begin{aligned} \mathcal{A}(\mathbf{d}) &= \left(\frac{2\pi m}{n} \right)^{-2 \sum_k \theta_k} \times \det \left\{ \frac{1}{m} \sum_{j=1}^m \operatorname{Re} [\Lambda_j(\boldsymbol{\theta})^{-1} \widehat{G}(\mathbf{d}_0) \overline{\Lambda_j(\boldsymbol{\theta})^{-1}}'] \right\} \\ &= \det \left\{ \frac{1}{m} \sum_{j=1}^m \operatorname{Re} [M_j(\boldsymbol{\theta}) \widehat{G}(\mathbf{d}_0) \overline{M_j(\boldsymbol{\theta})}'] \right\}, \end{aligned} \quad (\text{A.4})$$

where $M_j(\boldsymbol{\theta}) := \operatorname{diag}_{k \in \{1, \dots, q\}} \left\{ e^{i(\lambda_j - \pi)\theta_k/2} (j/m)^{\theta_k} \right\}$. To determine the function h in (i), we first show that

$$\frac{1}{m} \sum_{j=1}^m \operatorname{Re} [M_j(\boldsymbol{\theta}) \widehat{G}(\mathbf{d}_0) \overline{M_j(\boldsymbol{\theta})}'] = \frac{1}{m} \sum_{j=1}^m \operatorname{Re} [M_j(\boldsymbol{\theta}) G_0 \overline{M_j(\boldsymbol{\theta})}'] + o_{\mathbb{P}}(1), \quad (\text{A.5})$$

uniformly in Θ_1 . By Lemma 4.1, it follows that

$$\frac{1}{m} \sum_{j=1}^m \operatorname{Re} [M_j(\boldsymbol{\theta}) (\widehat{G}(\mathbf{d}_0) - G_0) \overline{M_j(\boldsymbol{\theta})}'] = \frac{1}{m} \sum_{j=1}^m \operatorname{Re} [M_j(\boldsymbol{\theta}) o_{\mathbb{P}}(1) \overline{M_j(\boldsymbol{\theta})}']. \quad (\text{A.6})$$

The (r, s) -th element in (A.6) is given by

$$\frac{1}{m} \sum_{j=1}^m \operatorname{Re} \left[e^{i(\lambda_j - \pi)(\theta_r - \theta_s)/2} \left(\frac{j}{m} \right)^{\theta_r + \theta_s} o_{\mathbb{P}}(1) \right] = \left[\frac{1}{m} \sum_{j=1}^m \left(\frac{j}{m} \right)^{\theta_r + \theta_s} \right] o_{\mathbb{P}}(1) = o_{\mathbb{P}}(1),$$

where the last equality follows from lemma 2 in Robinson (1995b) by taking $\gamma = \theta_r + \theta_s + 1 \in [2\epsilon, 1]$. Therefore, (A.5) follows uniformly in Θ_1 . The next step is to derive an approximation to the RHS of (A.5). First, since $e^{i(\lambda_j - \pi)(\theta_r - \theta_s)/2} = e^{i\pi(\theta_r - \theta_s)/2} + O(\lambda)$, the (r, s) -th element of the RHS of (A.5), omitting the $\operatorname{Re}[\cdot]$ operator, can be written as

$$\begin{aligned} \frac{1}{m} \sum_{j=1}^m e^{i(\lambda_j - \pi)(\theta_r - \theta_s)/2} \left(\frac{j}{m} \right)^{\theta_r + \theta_s} G_0^{rs} &= \\ &= e^{-i\pi(\theta_r - \theta_s)/2} \left[\frac{1}{m} \sum_{j=1}^m \left(\frac{j}{m} \right)^{\theta_r + \theta_s} G_0^{rs} \right] + \left[\frac{1}{m} \sum_{j=1}^m \left(\frac{j}{m} \right)^{\theta_r + \theta_s} G_0^{rs} \right] O\left(\frac{m}{n}\right). \end{aligned} \quad (\text{A.7})$$

Now, from lemma 2 in Robinson (1995b), it follows that

$$\begin{aligned} \frac{1}{m} \sum_{j=1}^m \left(\frac{j}{m} \right)^{\theta_r + \theta_s} &= \frac{1}{1 + \theta_r + \theta_s} \left(\frac{2\pi m}{n} \right)^{\theta_r + \theta_s} \left[\frac{1 + \theta_r + \theta_s}{m} \sum_{j=1}^m \left(\frac{j}{m} \right)^{\theta_r + \theta_s} \right] \\ &= \frac{1}{1 + \theta_r + \theta_s} \left(\frac{2\pi m}{n} \right)^{\theta_r + \theta_s} \left(1 + O(m^{-2\epsilon}) \right). \end{aligned} \quad (\text{A.8})$$

Substituting (A.8) on (A.7), it follows that

$$\frac{1}{m} \sum_{j=1}^m e^{i(\lambda_j - \pi)(\theta_r - \theta_s)/2} \left(\frac{j}{m} \right)^{\theta_r + \theta_s} o_{\mathbb{P}}(1) = \frac{e^{-i\pi(\theta_r - \theta_s)/2}}{1 + \theta_r + \theta_s} G_0^{rs} + O\left(\frac{m}{n}\right) + O(m^{-2\epsilon}). \quad (\text{A.9})$$

Define the matrices

$$\mathcal{E}(\boldsymbol{\theta}) := \left(e^{-i\pi(\theta_r - \theta_s)/2} \right)_{r,s=1}^q \quad \text{and} \quad \mathcal{M}(\boldsymbol{\theta}) := \left(\frac{1}{1 + \theta_r + \theta_s} \right)_{r,s=1}^q$$

and notice that, from (A.9), we have

$$\frac{1}{m} \sum_{j=1}^m M_j(\boldsymbol{\theta}) G_0 \overline{M_j(\boldsymbol{\theta})}' = \mathcal{E}(\boldsymbol{\theta}) \odot \mathcal{M}(\boldsymbol{\theta}) \odot G_0 + O\left(\frac{m}{n}\right) + O(m^{-2\epsilon}). \quad (\text{A.10})$$

Upon defining,

$$h(\mathbf{d}) := \det \left\{ \text{Re}[\mathcal{E}(\boldsymbol{\theta})] \odot \mathcal{M}(\boldsymbol{\theta}) \odot G_0 \right\},$$

the continuity of the determinant, the finiteness of $\mathcal{E}(\boldsymbol{\theta})$, $\mathcal{M}(\boldsymbol{\theta})$ and G_0 , for $\boldsymbol{\theta} \in \Theta_1$, and (A.10), imply (i). Inequality (ii) follows upon observing that $V(\boldsymbol{\theta}) := \text{Re}[\mathcal{E}(\boldsymbol{\theta})] \odot \mathcal{M}(\boldsymbol{\theta})$ is positive semidefinite (cf. Shimotsu, 2007, p.293) and, from Oppenheim's inequality,

$$h(\mathbf{d}) \geq \prod_{k=1}^q V_{kk}(\boldsymbol{\theta}) \det(G_0) = \prod_{k=1}^q \mathcal{M}_{kk}(\boldsymbol{\theta}) \det(G_0) = \mathcal{B}(\mathbf{d}).$$

By definition, $\mathcal{E}_{rs}(\mathbf{0}) = 1$, for all $r, s = 1, \dots, q$, so that

$$h(\mathbf{d}_0) = \det \left\{ \mathcal{M}(\mathbf{0}) \odot G_0 \right\} = \mathcal{B}(\mathbf{d}_0),$$

and (iii) follows.

Now we move to bound P_2 in (A.1). Expression (A.3) can be used to rewrite $L(\mathbf{d})$ as

$$\begin{aligned} L(\mathbf{d}) &= \log(\det\{\widehat{G}(\mathbf{d})\}) - \log(\det\{\widehat{G}(\mathbf{d}_0)\}) - 2 \sum_{k=1}^q \theta_k \frac{1}{m} \sum_{j=1}^m \log(\lambda_j) \\ &= \log(\det\{\widehat{\mathcal{D}}(\mathbf{d})\}) - \log(\det\{\widehat{\mathcal{D}}(\mathbf{d}_0)\}), \end{aligned} \quad (\text{A.11})$$

where

$$\widehat{\mathcal{D}}(\mathbf{d}) := \frac{1}{m} \sum_{j=1}^m \text{Re}[\mathcal{P}_j(\boldsymbol{\theta})^{-1} \widehat{G}(\mathbf{d}_0) \overline{\mathcal{P}_j(\boldsymbol{\theta})^{-1}}'],$$

with

$$\mathcal{P}_j(\boldsymbol{\theta}) := \text{diag}_{k \in \{1, \dots, q\}} \left\{ e^{i(\lambda_j - \pi)\theta_k/2} \left(\frac{j}{p}\right)^{\theta_k} \right\} \quad \text{and} \quad p := \exp\left(\frac{1}{m} \sum_{j=1}^m \log(j)\right),$$

and, as m tends to infinity, $p \sim m/e$. Observe that $\widehat{\mathcal{D}}(\mathbf{d})$ is positive semidefinite since each summand of $\widehat{\mathcal{D}}$ is. For $\kappa \in (0, 1)$, define

$$\widehat{\mathcal{D}}_\kappa(\mathbf{d}) := \frac{1}{m} \sum_{j=[m\kappa]}^m \text{Re}[\mathcal{P}_j(\boldsymbol{\theta})^{-1} \widehat{G}(\mathbf{d}_0) \overline{\mathcal{P}_j(\boldsymbol{\theta})^{-1}}'] \quad \text{and} \quad \mathcal{Q}_\kappa(\mathbf{d}) := \frac{1}{m} \sum_{j=[m\kappa]}^m \text{Re}[\mathcal{P}_j(\boldsymbol{\theta}) G_0 \overline{\mathcal{P}_j(\boldsymbol{\theta})}'],$$

where $[x]$ denotes the integer part of x . By Lemma 4.1,

$$\widehat{\mathcal{D}}_\kappa(\mathbf{d}) - \mathcal{Q}_\kappa(\mathbf{d}) = \frac{1}{m} \sum_{j=[m\kappa]}^m \text{Re}[\mathcal{P}_j(\boldsymbol{\theta})(\widehat{G}(\mathbf{d}_0) - G_0) \overline{\mathcal{P}_j(\boldsymbol{\theta})}'] = \frac{1}{m} \sum_{j=[m\kappa]}^m \text{Re}[\mathcal{P}_j(\boldsymbol{\theta}) o_{\mathbb{P}}(1) \overline{\mathcal{P}_j(\boldsymbol{\theta})}'].$$

The (r, s) -th element of $\widehat{\mathcal{D}}_\kappa(\mathbf{d}) - \mathcal{Q}_\kappa(\mathbf{d})$ is then

$$\begin{aligned} \left(\widehat{\mathcal{D}}_\kappa(\mathbf{d}) - \mathcal{Q}_\kappa(\mathbf{d})\right)_{rs} &= \text{Re} \left[\frac{1}{m} \sum_{j=[m\kappa]}^m e^{i(\lambda_j - \pi)(\theta_r - \theta_s)/2} \left(\frac{j}{p}\right)^{\theta_r + \theta_s} o_{\mathbb{P}}(1) \right] \\ &= \left(\frac{m}{p}\right)^{\theta_r + \theta_s} \frac{1}{m} \sum_{j=[m\kappa]}^m \left(\frac{j}{m}\right)^{\theta_r + \theta_s} o_{\mathbb{P}}(1) \\ &= \left(\frac{m}{p}\right)^{\theta_r + \theta_s} O(1) o_{\mathbb{P}}(1) = o_{\mathbb{P}}(1), \end{aligned}$$

uniformly in $\boldsymbol{\theta} \in \Theta_2$, where the penultimate equality follows from lemma 5.4 in Shimotsu and Philips (2005). Therefore, as $n \rightarrow \infty$, for any $\kappa \in (0, 1)$,

$$\sup_{\Theta_2} \left\{ \left| \det\{\widehat{\mathcal{D}}_\kappa(\mathbf{d})\} - \det\{\mathcal{Q}_\kappa(\mathbf{d})\} \right| \right\} = o_{\mathbb{P}}(1).$$

Now, for all $\mathbf{d} \in \Theta_2$, one can obtain $\Delta \in (0, 0.1)$ so that, for κ sufficiently small and m sufficiently large (cf. Shimotsu, 2007, p.295),

$$\inf_{\Theta_2} \{ \det \{ \widehat{\mathcal{D}}_\kappa(\mathbf{d}) \} \} = \inf_{\Theta_2} \{ \det \{ \mathcal{Q}_\kappa(\mathbf{d}) \} \} + o_{\mathbb{P}}(1) \geq \det\{G_0\}(1 + \Delta) + o_{\mathbb{P}}(1).$$

Now, since $\det \{ \widehat{\mathcal{D}}(\mathbf{d}_0) \} = \det \{ \widehat{G}(\mathbf{d}_0) \} \xrightarrow{\mathbb{P}} \det\{G_0\}$, as $n \rightarrow \infty$, by Lemma 4.1, it follows that

$$\mathbb{P} \left(\inf_{\Theta_2} \{ \det \{ \widehat{\mathcal{D}}_\kappa(\mathbf{d}) \} \} - \det \{ \widehat{\mathcal{D}}(\mathbf{d}_0) \} \leq 0 \right) \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Also, since $\det \{ \widehat{\mathcal{D}}(\mathbf{d}) \} \geq \det \{ \widehat{\mathcal{D}}_\kappa(\mathbf{d}) \}$ we conclude that

$$\mathbb{P} \left(\inf_{\Theta_2} \{ \det \{ \widehat{\mathcal{D}}(\mathbf{d}) \} \} - \det \{ \widehat{\mathcal{D}}(\mathbf{d}_0) \} \leq 0 \right) \rightarrow 0, \quad \text{as } n \rightarrow \infty. \quad (\text{A.12})$$

Since the logarithm is a monotone increasing function of its argument and in view of (A.11), (A.12) implies $P_2 \xrightarrow{n \rightarrow \infty} 0$ and the proof is complete. \blacksquare

Proof of Lemma 4.2:

For fixed $r, s \in \{1, \dots, q\}$, let $\mathcal{A}_{uv} := \sum_{j=u}^v \mathcal{A}_j$ and $\mathcal{B}_{uv} := \sum_{j=u}^v \mathcal{B}_j$, where

$$\mathcal{A}_j := e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} [f_n^{rs}(\lambda_j) - (A(\lambda_j))_{r \cdot} I_\epsilon(\lambda_j) (\overline{A(\lambda_j)})'_{\cdot s}], \quad (\text{A.13})$$

and

$$\mathcal{B}_j := e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} (A(\lambda_j))_{r \cdot} I_\epsilon(\lambda_j) (\overline{A(\lambda_j)})'_{\cdot s} - G_0^{rs}. \quad (\text{A.14})$$

Hence, for each j , $\mathcal{A}_j + \mathcal{B}_j = e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} f_n^{rs}(\lambda_j) - G_0^{rs}$. For fixed $u \leq j \leq v$, we have

$$\begin{aligned} \mathbb{E}(|\mathcal{A}_j|) &= (2\pi j)^{d_r^0 + d_s^0} \mathbb{E} \left(n^{-d_r^0 - d_s^0} \left| f_n^{rs}(\lambda_j) - (A(\lambda_j))_{r \cdot} I_\epsilon(\lambda_j) (\overline{A(\lambda_j)})'_{\cdot s} \right| \right) \\ &= (2\pi j)^{d_r^0 + d_s^0} O \left(\frac{1}{j^{d_r^0 + d_s^0 + \gamma}} \right) = O(j^{-\gamma}). \end{aligned}$$

Therefore, $\max_{r,s} \{ \mathbb{E}(|\sum_{j=u}^v \mathcal{A}_j|) \} = O(v^{1-\gamma})$ and the result on \mathcal{A}_{uv} follows. As for \mathcal{B}_j , from the proof of lemma 1(a) in Shimotsu (2007) (notice that \mathcal{B}_j does not depend on f_n) it follows that $\sum_{j=u}^v \mathcal{B}_j = o_{\mathbb{P}}(v)$ uniformly in u and v and hence the desired result on \mathcal{B}_{uv} follows. \blacksquare

Proof of Theorem 4.2:

From a careful inspection of the proof of Theorem 4.1, we observe that it suffices to show (with the same notation as in that proof)

$$\frac{1}{m} \sum_{j=1}^m \text{Re} [M_j(\boldsymbol{\theta}) \widehat{G}(\mathbf{d}_0) \overline{M_j(\boldsymbol{\theta})}]' = \frac{1}{m} \sum_{j=1}^m \text{Re} [M_j(\boldsymbol{\theta}) G_0 \overline{M_j(\boldsymbol{\theta})}]' + o_{\mathbb{P}}(1), \quad (\text{A.15})$$

uniformly in Θ_1 and that $\widehat{\mathcal{D}}_\kappa(\mathbf{d}) - \mathcal{Q}_\kappa(\mathbf{d}) = o_{\mathbb{P}}(1)$, uniformly in Θ_2 . To show (A.15), notice that the (r, s) -th component of the LHS in (A.15) is given by

$$\frac{1}{m} \sum_{j=1}^m \text{Re} \left[e^{i(\lambda_j - \pi)(\theta_r - \theta_s)/2} \left(\frac{j}{m} \right)^{\theta_r + \theta_s} f_n^{rs}(\lambda_j) \left(\Lambda_j^{(r)}(\mathbf{d}_0) \overline{\Lambda_j^{(s)}(\mathbf{d}_0)} \right)' \right]^{-1}.$$

Summation by parts (see Zygmund, 2002, p.3) yields

$$\sup_{\Theta_1} \left\{ \left| \frac{1}{m} \sum_{j=1}^m e^{i(\lambda_j - \pi)(\theta_r - \theta_s)/2} \left(\frac{j}{m} \right)^{\theta_r + \theta_s} \left[f_n^{rs}(\lambda_j) \left(\Lambda_j^{(r)}(\mathbf{d}_0) \overline{\Lambda_j^{(s)}(\mathbf{d}_0)} \right)' - G_0^{rs} \right] \right| \right\} \leq$$

$$\begin{aligned}
&\leq \frac{1}{m} \sum_{k=1}^{m-1} \sup_{\Theta_1} \left\{ \left| e^{i(\lambda_k - \pi)(\theta_r - \theta_s)/2} \left(\frac{k}{m} \right)^{\theta_r + \theta_s} - e^{i(\lambda_{k+1} - \pi)(\theta_r - \theta_s)/2} \left(\frac{k+1}{m} \right)^{\theta_r + \theta_s} \right| \right\} \times \\
&\quad \times \left| \sum_{j=1}^k \left[f_n^{rs}(\lambda_j) \left(\Lambda_j^{(r)}(\mathbf{d}_0) \overline{\Lambda_j^{(s)}(\mathbf{d}_0)} \right)'^{-1} - G_0^{rs} \right] \right| + \left| \frac{1}{m} \sum_{j=1}^m \left[f_n^{rs}(\lambda_j) \left(\Lambda_j^{(r)}(\mathbf{d}_0) \overline{\Lambda_j^{(s)}(\mathbf{d}_0)} \right)'^{-1} - G_0^{rs} \right] \right| \\
&\leq C \sum_{k=1}^{m-1} \left(\frac{k}{m} \right)^{2\epsilon} \frac{1}{k^2} \left| \sum_{j=1}^k \left[f_n^{rs}(\lambda_j) \left(\Lambda_j^{(r)}(\mathbf{d}_0) \overline{\Lambda_j^{(s)}(\mathbf{d}_0)} \right)'^{-1} - G_0^{rs} \right] \right| + \\
&\quad + \left| \frac{1}{m} \sum_{j=1}^m \left[f_n^{rs}(\lambda_j) \left(\Lambda_j^{(r)}(\mathbf{d}_0) \overline{\Lambda_j^{(s)}(\mathbf{d}_0)} \right)'^{-1} - G_0^{rs} \right] \right|, \tag{A.16}
\end{aligned}$$

where $0 < C < \infty$ is a constant. Now, from Lemma 4.2,

$$\begin{aligned}
&\sum_{k=1}^{m-1} \left(\frac{k}{m} \right)^{2\epsilon} \frac{1}{k^2} \left| \sum_{j=1}^k \left[f_n^{rs}(\lambda_j) \left(\Lambda_j^{(r)}(\mathbf{d}_0) \overline{\Lambda_j^{(s)}(\mathbf{d}_0)} \right)'^{-1} - G_0^{rs} \right] \right| \leq \\
&\leq \sum_{k=1}^{m-1} \left(\frac{k}{m} \right)^{2\epsilon} \frac{1}{k^2} (|\mathcal{A}_{1k}| + |\mathcal{B}_{1k}|) \\
&= \frac{1}{m^{2\epsilon}} \sum_{k=1}^{m-1} k^{2(\epsilon-1)} |\mathcal{A}_{1k}| + \frac{1}{m^{2\epsilon}} \sum_{k=1}^{m-1} k^{2(\epsilon-1)} o_{\mathbb{P}}(k) \\
&= o_{\mathbb{P}}(1) + \frac{m(m-1)^{2\epsilon-1}}{m^{2\epsilon}} o_{\mathbb{P}}(1) = o_{\mathbb{P}}(1),
\end{aligned}$$

uniformly in (r, s) , where the penultimate equality follows from Lemma 4.2 which implies

$$\begin{aligned}
\mathbb{E} \left(\frac{1}{m^{2\epsilon}} \sum_{k=1}^{m-1} k^{2(\epsilon-1)} |\mathcal{A}_{1k}| \right) &= \frac{1}{m^{2\epsilon}} \sum_{k=1}^{m-1} k^{2(\epsilon-1)} \mathbb{E}(|\mathcal{A}_{1k}|) \\
&\leq \frac{(m-1)^{2\epsilon-1}}{m^{2\epsilon}} O((m-1)^{1-\gamma}) = o(1),
\end{aligned}$$

uniformly in (r, s) . The other term in (A.16) is also $o_{\mathbb{P}}(1)$, uniformly in (r, s) , by the same argument and, hence, (A.15) follows. On the other hand, the (r, s) -th element of $\widehat{\mathcal{D}}_{\kappa}(\mathbf{d}) - \mathcal{Q}_{\kappa}(\mathbf{d})$ is given by

$$\begin{aligned}
&\frac{1}{m} \sum_{j=[m\kappa]}^m \operatorname{Re} \left[e^{i(\lambda_j - \pi)(\theta_r - \theta_s)/2} \left(\frac{j}{p} \right)^{\theta_r + \theta_s} \left[f_n^{rs}(\lambda_j) \left(\Lambda_j^{(r)}(\mathbf{d}_0) \overline{\Lambda_j^{(s)}(\mathbf{d}_0)} \right)'^{-1} - G_0^{rs} \right] \right] = \\
&= \left(\frac{m}{p} \right)^{\theta_r + \theta_s} \operatorname{Re} \left[\frac{1}{m} \sum_{j=[m\kappa]}^m e^{i(\lambda_j - \pi)(\theta_r - \theta_s)/2} \left(\frac{j}{m} \right)^{\theta_r + \theta_s} \left[f_n^{rs}(\lambda_j) \left(\Lambda_j^{(r)}(\mathbf{d}_0) \overline{\Lambda_j^{(s)}(\mathbf{d}_0)} \right)'^{-1} - G_0^{rs} \right] \right] \\
&= o_{\mathbb{P}}(1),
\end{aligned}$$

uniformly in $\boldsymbol{\theta} \in \Theta_2$, where the last equality is derived similarly to (A.16) from summation by parts and lemma 5.4 in Shimotsu and Phillips (2005). This completes the proof. \blacksquare

Proof of Corollary 4.1:

From the proof of Lemma 4.2, in order to show (4.3) it suffices to show that, for \mathcal{A}_j as in (A.13) and f_n as in the enunciate, $\mathbb{E}(|\mathcal{A}_j|) = O(j^{-\gamma})$, for some $\gamma > 0$. From the proof of lemma 1(a) in Shimotsu (2007) (see also theorem 2 in Robinson, 1995a), we have

$$\mathbb{E}(I_n(\lambda_j)) = f(\lambda_j) \left(1 + O \left(\frac{\log(j+1)}{j} \right) \right); \tag{A.17a}$$

$$\mathbb{E}(I_{\epsilon}(\lambda_j)) = \frac{I_q}{2\pi} + O \left(\frac{\log(j+1)}{j} \right); \tag{A.17b}$$

$$\mathbb{E}(w_n^r(\lambda_j) \overline{w_\varepsilon^s(\lambda_j)})' = \frac{(A(\lambda_j))_{r\cdot}}{2\pi} + O\left(\frac{\log(j+1)}{j\lambda_j^{d_r^0}}\right), \quad \text{for } j = 1, \dots, m. \quad (\text{A.17c})$$

Assumption **A3** implies

$$\begin{aligned} \mathbb{E}(|\mathcal{A}_j|) &= \mathbb{E}\left(\left|\sum_{|k| \leq \ell(n)} W_n^{rs}(k) \left(e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} [I_n^{rs}(\lambda_{j+k}) - (A(\lambda_j))_{r\cdot} I_\varepsilon(\lambda_j) \overline{(A(\lambda_j))'_{\cdot s}}]\right)\right|\right) \\ &\leq \sum_{|k| \leq \ell(n)} W_n^{rs}(k) \mathbb{E}\left(\left|e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} [I_n^{rs}(\lambda_{j+k}) - (A(\lambda_j))_{r\cdot} I_\varepsilon(\lambda_j) \overline{(A(\lambda_j))'_{\cdot s}}]\right|\right). \end{aligned} \quad (\text{A.18})$$

Rewrite the expression inside the squared brackets on the RHS of (A.18) as

$$\left[(w_n^r(\lambda_{j+k}) - (A(\lambda_j))_{r\cdot} w_\varepsilon(\lambda_j)) \overline{w_n^s(\lambda_{j+k})}' + (A(\lambda_j))_{r\cdot} w_\varepsilon(\lambda_j) \left[\overline{w_n^s(\lambda_{j+k})}' - \overline{(A(\lambda_j))'_{\cdot s} w_\varepsilon(\lambda_j)}\right]\right].$$

Now, observe that

$$\begin{aligned} w_n^r(\lambda_{j+k}) \overline{w_\varepsilon^s(\lambda_j)}' &= \frac{1}{2\pi n} \left(\sum_{t=1}^n \mathbf{X}_t^{(r)} e^{it\lambda_{j+k}}\right) \left(\sum_{t=1}^n \varepsilon_t^{(s)} e^{-it\lambda_j}\right)' = \frac{1}{2\pi n} \sum_{a=1}^n \sum_{b=1}^n \mathbf{X}_a^{(r)} \varepsilon_b^{(s)'} e^{i(a-b)\lambda_j} e^{ia\lambda_k} \\ &= \frac{1}{2\pi n} \sum_{a=1}^n \sum_{b=1}^n \left(\sum_{t=0}^\infty A_t^r \varepsilon_{a-t}^{(r)}\right) \varepsilon_b^{(s)'} e^{i(a-b)\lambda_j} e^{ia\lambda_k} = \frac{1}{2\pi n} \sum_{t=0}^\infty A_t^r e^{it\lambda_j} I_\varepsilon^{rs}(\lambda_j) O(1), \end{aligned}$$

so that, by using (A.17b), we conclude that

$$\mathbb{E}\left(w_n^r(\lambda_{j+k}) \overline{w_\varepsilon^s(\lambda_j)}'\right) = \frac{(A(\lambda_j))_{r\cdot}}{2\pi} O(1) + O\left(\frac{\log(j+1)}{j\lambda_j^{d_r^0}}\right). \quad (\text{A.19})$$

From (A.17a), (A.17b), (A.17c), $(A(\lambda_j))_{r\cdot} \overline{(A(\lambda_j))'_{\cdot s}} / 2\pi = f^{rr}(\lambda_j)$, $f^{rr}(\lambda_j)^{2d_r^0} \sim G_0^{rr}$, assumption **A5** and because the function $\log(|x| + 1)/|x|$ is increasing in $(-\infty, 0)$ and decreasing in $(0, \infty)$, it follows that

$$\begin{aligned} \mathbb{E}\left(\left|w_n^r(\lambda_{j+k}) - (A(\lambda_j))_{r\cdot} w_\varepsilon(\lambda_j)\right|^2\right) &= \mathbb{E}\left(\left|I_n^{rr}(\lambda_{j+k}) - w_n^r(\lambda_{j+k}) \overline{w_\varepsilon(\lambda_j)}' \overline{(A(\lambda_j))'_{\cdot s}}\right.\right. \\ &\quad \left.\left.+ (A(\lambda_j))_{r\cdot} I_\varepsilon(\lambda_j) \overline{(A(\lambda_j))'_{\cdot s}} - \overline{(A(\lambda_j))'_{\cdot s}} w_\varepsilon(\lambda_j) w_n^r(\lambda_{j+k})\right|^2\right) \\ &= O\left(\frac{\log(\min\{|j+k|, j\} + 1)}{\min\{|j+k|, j\} \lambda_{j+k}^{2d_r^0}}\right) = O(\psi(j, k, d_r^0)), \end{aligned}$$

where $\psi(j, k, d_r^0) := \frac{\log(\min\{|j+k|, j\} + 1)}{\min\{|j+k|, j\} \lambda_{j+k}^{2d_r^0}}$, and similarly for $\mathbb{E}(\left|\overline{w_n^s(\lambda_j)}' - \overline{(A(\lambda_j))'_{\cdot s} w_\varepsilon(\lambda_j)}\right|^2)$. Also,

$\mathbb{E}(I_n^{ss}(\lambda_{j+k})) = O(\lambda_{j+k}^{-2d_s^0})$. By the Cauchy-Schwartz inequality,

$$\begin{aligned} \mathbb{E}(|\mathcal{A}_j|) &\leq \sum_{|k| \leq \ell(n)} W_n^{rs}(k) |\lambda_j^{d_r^0 + d_s^0}| \left[\mathbb{E}\left(\left|w_n^r(\lambda_j) - (A(\lambda_j))_{r\cdot} w_\varepsilon(\lambda_j)\right|^2\right)^{\frac{1}{2}} \mathbb{E}\left(\left|w_n^r(\lambda_{j+k})\right|^2\right)^{\frac{1}{2}} \right. \\ &\quad \left. + \mathbb{E}\left(\left|(A(\lambda_j))_{r\cdot} w_\varepsilon(\lambda_j)\right|^2\right)^{\frac{1}{2}} \mathbb{E}\left(\left|\overline{w_n^s(\lambda_j)}' - \overline{(A(\lambda_j))'_{\cdot s} w_\varepsilon(\lambda_j)}\right|^2\right)^{\frac{1}{2}} \right] \\ &= \sum_{|k| \leq \ell(n)} W_n^{rs}(k) |\lambda_j|^{d_r^0 + d_s^0} O(\sqrt{\psi(j, k, d_r^0)}) O(|\lambda_{j+k}|^{-d_s^0}) \\ &= \sum_{|k| \leq \ell(n)} W_n^{rs}(k) O\left(\frac{\log(\min\{|j+k|, j\} + 1)^{1/2}}{\min\{|j+k|, j\}^{1/2}}\right) O\left(\left|\frac{\lambda_j}{\lambda_{j+k}}\right|^{d_r^0 + d_s^0}\right) = O(j^{-\gamma_0}), \end{aligned}$$

for $0 < \gamma_0 < 1/2$, where the last equality follows from assumption **A5**. This concludes the proof. ■

Proof of Lemma 5.1:

(a) For $r, s \in \{1, \dots, q\}$ fixed, ignoring the maximum in expression (5.2) for a while, its argument can be written as

$$\begin{aligned} \left| \sum_{j=1}^v e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} \left[f_n^{rs}(\lambda_j) - (A(\lambda_j))_{r.} I_{\epsilon}(\lambda_j) (\overline{A(\lambda_j)})'_{.s} \right] \right| &\leq \\ &\leq \sum_{j=1}^v \left(\frac{2\pi j}{n} \right)^{d_r^0 + d_s^0} \left| f_n^{rs}(\lambda_j) - (A(\lambda_j))_{r.} I_{\epsilon}(\lambda_j) (\overline{A(\lambda_j)})'_{.s} \right| \\ &\leq O\left(\frac{v^{d_r^0 + d_s^0}}{n^{d_r^0 + d_s^0}} \right) o_{\mathbb{P}}\left(\frac{n^{d_r^0 + d_s^0}}{\log(v) v^{d_r^0 + d_s^0 - 1/2}} \right) = o_{\mathbb{P}}\left(\frac{v^{1/2}}{\log(v)} \right), \end{aligned}$$

uniformly in $1 \leq v \leq m$ and the result follows.

(b) Rewrite the argument of the summation in (5.3) as $\mathcal{A}_j + \mathcal{B}_j + \mathcal{C}_j$, where

$$\begin{aligned} \mathcal{A}_j &:= e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} \left[f_n^{rs}(\lambda_j) - (A(\lambda_j))_{r.} I_{\epsilon}(\lambda_j) (\overline{A(\lambda_j)})'_{.s} \right], \\ \mathcal{B}_j &:= e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} \left[(A(\lambda_j))_{r.} I_{\epsilon}(\lambda_j) (\overline{A(\lambda_j)})'_{.s} - f_{rs}(\lambda_j) \right], \\ \mathcal{C}_j &:= e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} f_{rs}(\lambda_j) - G_0^{rs}. \end{aligned}$$

Part (a) yields $\max_{r,s} \left\{ \sum_{j=1}^v |\mathcal{A}_j| \right\} = o_{\mathbb{P}}\left(v^{1/2} (\log(v))^{-1} \right)$, while, from the proof of lemma 1(b2) in Shimotsu (2007), we obtain $\max_{r,s} \left\{ \sum_{j=1}^v |\mathcal{B}_j| \right\} = O_{\mathbb{P}}(v^{1/2} \log(v))$. Assumption **C1** implies $\max_{r,s} \left\{ \sum_{j=1}^v |\mathcal{C}_j| \right\} = O(v^{\alpha+1} n^{-\alpha})$. The result now follows by noticing that $v^{1/2} (\log(v))^{-1} = O(v^{1/2} \log(v))$. \blacksquare

Proof of Theorem 5.1:

The idea of the proof is similar to that of Lobato (1999) with similar adaptations as in Shimotsu (2007). By hypothesis,

$$\mathbf{0} = \frac{\partial S(\mathbf{d})}{\partial \mathbf{d}} \Big|_{\hat{\mathbf{d}}} = \frac{\partial S(\mathbf{d})}{\partial \mathbf{d}} \Big|_{\mathbf{d}_0} + \left(\frac{\partial^2 S(\mathbf{d})}{\partial \mathbf{d} \partial \mathbf{d}'} \Big|_{\bar{\mathbf{d}}} \right) (\hat{\mathbf{d}} - \mathbf{d}_0),$$

with probability tending to 1, as n tends to infinity, for some $\bar{\mathbf{d}}$ such that $\|\bar{\mathbf{d}} - \mathbf{d}_0\|_{\infty} \leq \|\hat{\mathbf{d}} - \mathbf{d}_0\|_{\infty}$. We observe that $\hat{\mathbf{d}}$ has the stated limiting distribution if

$$\frac{\partial S(\mathbf{d})}{\partial \mathbf{d}} \Big|_{\mathbf{d}_0} \xrightarrow{d} N(0, \Omega) \tag{A.20}$$

and

$$\frac{\partial^2 S(\mathbf{d})}{\partial \mathbf{d} \partial \mathbf{d}'} \Big|_{\bar{\mathbf{d}}} \xrightarrow{\mathbb{P}} \Omega. \tag{A.21}$$

We shall prove (A.20) first. In order to do that, we use a Crámer-Wold device. Let η be an arbitrary vector in \mathbb{R}^q . Observe that, for $r \in \{1, \dots, q\}$,

$$\sqrt{m} \frac{\partial S(\mathbf{d})}{\partial d_r} = -\frac{2}{\sqrt{m}} \sum_{j=1}^m \log(\lambda_j) + \text{tr} \left[\hat{G}(\mathbf{d})^{-1} \sqrt{m} \frac{\partial \hat{G}(\mathbf{d})}{\partial d_r} \right].$$

Let $\mathbf{I}_{(r)}$ denote a $q \times q$ matrix whose (r, r) -th element is 1 and all other elements are zero. Define a function $\varphi : (0, \infty) \rightarrow \mathbb{C}$ by

$$\varphi(x) := \log(x) + i \left(\frac{x - \pi}{2} \right). \tag{A.22}$$

Since $\Lambda_j(\mathbf{d})^{-1} = \text{diag}_{k \in \{1, \dots, q\}} \{\lambda_j^{d_k} e^{i(\lambda_j - \pi)d_k/2}\}$ and $\text{Re}[(a + ib)(c + id)] = ac - bd$, we can write

$$\begin{aligned} \sqrt{m} \frac{\partial \widehat{G}(\mathbf{d})}{\partial d_r} \Big|_{\mathbf{d}_0} &= \frac{1}{\sqrt{m}} \sum_{j=1}^m \text{Re} \left[\varphi(\lambda_j) \Lambda_j(\mathbf{d}_0)^{-1} \mathbf{I}_{(r)} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] + \\ &\quad + \frac{1}{\sqrt{m}} \sum_{j=1}^m \text{Re} \left[\overline{\varphi(\lambda_j)} \Lambda_j(\mathbf{d}_0)^{-1} f_n(\lambda_j) \mathbf{I}_{(r)} \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] \\ &= \frac{1}{\sqrt{m}} \sum_{j=1}^m \log(\lambda_j) \text{Re} \left[\Lambda_j(\mathbf{d}_0)^{-1} (\mathbf{I}_{(r)} f_n(\lambda_j) + f_n(\lambda_j) \mathbf{I}_{(r)}) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] + \\ &\quad + \frac{1}{\sqrt{m}} \sum_{j=1}^m \left[\frac{\lambda_j - \pi}{2} \right] \text{Im} \left[\Lambda_j(\mathbf{d}_0)^{-1} (-\mathbf{I}_{(r)} f_n(\lambda_j) + f_n(\lambda_j) \mathbf{I}_{(r)}) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right], \\ &:= \mathcal{H}_1(r) + \mathcal{H}_2(r). \end{aligned} \tag{A.23}$$

Therefore, from (A.23), we obtain

$$\begin{aligned} \eta' \sqrt{m} \frac{\partial S(\mathbf{d})}{\partial \mathbf{d}} \Big|_{\mathbf{d}_0} &= \sum_{k=1}^q \eta_k \sqrt{m} \frac{\partial S(\mathbf{d})}{\partial d_k} \Big|_{\mathbf{d}_0} = \\ &= \sum_{k=1}^q \eta_k \left[-\frac{2}{\sqrt{m}} \sum_{j=1}^m \log(\lambda_j) + \text{tr} [\widehat{G}(\mathbf{d}_0)^{-1} \mathcal{H}_1(k)] \right] + \sum_{k=1}^q \eta_k \text{tr} [\widehat{G}(\mathbf{d}_0)^{-1} \mathcal{H}_2(k)], \\ &:= \mathcal{R}_1 + \mathcal{R}_2. \end{aligned}$$

We analyze \mathcal{R}_1 first. By letting

$$a_j := \log(\lambda_j) - \frac{1}{m} \sum_{k=1}^m \log(\lambda_k) = \log(j) - \frac{1}{m} \sum_{k=1}^m \log(k) = O(\log(m)),$$

we can write

$$\begin{aligned} -\frac{2}{\sqrt{m}} \sum_{j=1}^m \log(\lambda_j) + \text{tr} [\widehat{G}(\mathbf{d}_0)^{-1} \mathcal{H}_1(k)] &= \text{tr} \left[\widehat{G}(\mathbf{d}_0)^{-1} \left(\mathcal{H}_1(k) - \frac{2}{\sqrt{m}} \sum_{j=1}^m \log(\lambda_j) \widehat{G}(\mathbf{d}_0) \mathbf{I}_{(k)} \right) \right] \\ &= \text{tr} \left[\widehat{G}(\mathbf{d}_0)^{-1} \frac{2}{\sqrt{m}} \sum_{j=1}^m a_j \text{Re} \left[\Lambda_j(\mathbf{d}_0)^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] \mathbf{I}_{(k)} \right]. \end{aligned} \tag{A.24}$$

By Lemma 5.1(b), (A.24) can be written as

$$\left[(G_0^{-1})_{k \cdot} + o_{\mathbb{P}}(1) \right] \frac{2}{\sqrt{m}} \sum_{j=1}^m a_j \left(\text{Re} \left[\Lambda_j(\mathbf{d}_0)^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] \right)_{\cdot k}. \tag{A.25}$$

Now, by Lemma 5.1(a),

$$\begin{aligned} &\left| \left(\sum_{j=1}^m a_j \Lambda_j(\mathbf{d}_0)^{-1} \left(f_n(\lambda_j) - A(\lambda_j) I_{\varepsilon}(\lambda_j) \overline{A(\lambda_j)}' \right) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right)_{rs} \right| \leq \\ &\leq O(\log(m)) \max_{v=1, \dots, m} \left\{ \left| \sum_{j=1}^v e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} \left(f_n^{rs}(\lambda_j) - (A(\lambda_j))_{r \cdot} I_{\varepsilon}(\lambda_j) \overline{A(\lambda_j)}'_{\cdot s} \right) \right| \right\} \\ &= O(\log(m)) o_{\mathbb{P}} \left(\frac{\sqrt{m}}{\log(m)} \right) = o_{\mathbb{P}}(\sqrt{m}), \end{aligned}$$

uniformly in $r, s \in \{1, \dots, q\}$. Therefore,

$$\frac{1}{\sqrt{m}} \sum_{j=1}^m a_j \Lambda_j(\mathbf{d}_0)^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' =$$

$$\begin{aligned}
&= \frac{1}{\sqrt{m}} \sum_{j=1}^m a_j \Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) I_{\varepsilon}(\lambda_j) \overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' + \frac{1}{\sqrt{m}} o_{\mathbb{P}}(\sqrt{m}) \\
&= \frac{1}{\sqrt{m}} \sum_{j=1}^m a_j \left[\Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) I_{\varepsilon}(\lambda_j) \overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' - G_0 \right] + o_{\mathbb{P}}(1), \tag{A.26}
\end{aligned}$$

where the last equality follows from $\sum_{j=1}^m a_j = 0$. For the moment, ignore the $o_{\mathbb{P}}(1)$ term in (A.25), which will be dealt with later. By doing so and applying (A.26), it follows that

$$\mathcal{R}_1 = \frac{2}{\sqrt{m}} \sum_{k=1}^q \eta_k \sum_{j=1}^m a_j \left[(G_0^{-1})_{k \cdot} \left(\operatorname{Re} \left[\Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) I_{\varepsilon}(\lambda_j) \overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] \right)_{\cdot k} - 1 \right] + o_{\mathbb{P}}(1).$$

Writing

$$I_{\varepsilon}(\lambda_j) = \frac{1}{2\pi n} \sum_{t=1}^n \varepsilon_t \varepsilon'_t + \frac{1}{2\pi n} \sum_{\substack{u,v=1 \\ u \neq v}}^n \varepsilon_u \varepsilon'_v e^{i(u-v)\lambda_j} := J_1(n) + J_2(j; n), \tag{A.27}$$

we can rewrite \mathcal{R}_1 as

$$\begin{aligned}
\mathcal{R}_1 &= \frac{2}{\sqrt{m}} \sum_{k=1}^q \eta_k \sum_{j=1}^m a_j \left[(G_0^{-1})_{k \cdot} \left(\operatorname{Re} \left[\Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) J_1(n) \overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] \right)_{\cdot k} - 1 \right] + \\
&+ \frac{2}{\sqrt{m}} \sum_{k=1}^q \eta_k \sum_{j=1}^m a_j \left[(G_0^{-1})_{k \cdot} \left(\operatorname{Re} \left[\Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) J_2(j; n) \overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] \right)_{\cdot k} \right] + o_{\mathbb{P}}(1). \tag{A.28}
\end{aligned}$$

The first part of the RHS in (A.28) is $o_{\mathbb{P}}(1)$, which follows (cf. Shimotsu, 2007, p.297) from applying the proof of lemma 2 in appendix D of Lobato (1999), so that we can write

$$\mathcal{R}_1 = \sum_{u=2}^n \varepsilon'_u \sum_{v=1}^{t-1} \zeta_{u-v} \varepsilon_v + o_{\mathbb{P}}(1) \quad \text{with} \quad \zeta_t := \frac{1}{\pi n \sqrt{m}} \sum_{j=1}^m a_j \operatorname{Re} \left[\Omega_j e^{-it\lambda_j} + \Omega'_j e^{it\lambda_j} \right],$$

where

$$\Omega_j := \sum_{k=1}^q \eta_k \left(\overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right)_{\cdot k} (G_0^{-1})_{k \cdot} \Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j).$$

By writing

$$\xi_t := \frac{1}{\pi n \sqrt{m}} \sum_{j=1}^m a_j \operatorname{Re} [\Omega_j + \Omega'_j] \cos(t\lambda_j),$$

\mathcal{R}_1 can be further simplified, by using the argument in the proof of theorem 2, p.297, in Shimotsu (2007), as

$$\mathcal{R}_1 = \sum_{u=2}^n \varepsilon'_u \sum_{v=1}^{t-1} \xi_{u-v} \varepsilon_v + o_{\mathbb{P}}(1). \tag{A.29}$$

As for \mathcal{R}_2 , proceeding analogously to \mathcal{R}_1 , we obtain

$$\operatorname{tr} \left[\widehat{G}(\mathbf{d}_0)^{-1} \mathcal{H}_2(k) \right] = \left[(G_0^{-1})_{k \cdot} + o_{\mathbb{P}}(1) \right] \frac{1}{\sqrt{m}} \sum_{j=1}^m (\lambda_j - \pi) \left(\operatorname{Im} \left[\Lambda_j(\mathbf{d}_0)^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] \right)_{\cdot k}. \tag{A.30}$$

On one hand,

$$\begin{aligned}
&\left(\sum_{j=1}^m (\lambda_j - \pi) \Lambda_j(\mathbf{d}_0)^{-1} \left(f_n(\lambda_j) - A(\lambda_j) I_{\varepsilon}(\lambda_j) \overline{A(\lambda_j)}' \right) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right)_{rs} = \\
&= \sum_{j=1}^m (\lambda_j - \pi) e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} \left(f_n^{rs}(\lambda_j) - (A(\lambda_j))_{r \cdot} I_{\varepsilon}(\lambda_j) (\overline{A(\lambda_j)})'_{\cdot s} \right) \\
&\leq \pi \left(\frac{m}{n} - 1 \right) \sum_{j=1}^m e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} \left(f_n^{rs}(\lambda_j) - (A(\lambda_j))_{r \cdot} I_{\varepsilon}(\lambda_j) (\overline{A(\lambda_j)})'_{\cdot s} \right)
\end{aligned}$$

$$= O(1) o_{\mathbb{P}}\left(\frac{\sqrt{m}}{\log(m)}\right) = o_{\mathbb{P}}\left(\frac{\sqrt{m}}{\log(m)}\right). \quad (\text{A.31})$$

On the other hand, by using (A.27),

$$\begin{aligned} \sum_{j=1}^m \lambda_j \Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) I_{\epsilon}(\lambda_j) \overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' &\leq \\ &\leq \frac{2\pi m}{n} \sum_{j=1}^m \Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) I_{\epsilon}(\lambda_j) \overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \\ &= \frac{m}{n} \sum_{j=1}^m \Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) \left(\frac{1}{n} \sum_{t=1}^n \epsilon_t \epsilon_t' \right) \overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' + \\ &\quad + \frac{m}{n} \sum_{j=1}^m \Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) \left(\frac{1}{n} \sum_{\substack{u,v=1 \\ u \neq v}}^n \epsilon_u \epsilon_v' e^{i(u-v)\lambda_j} \right) \overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \\ &:= \mathcal{S}_1 + \mathcal{S}_2. \end{aligned}$$

From theorem 1 in Heyde and Senata (1972) and assumption **C2**, $\frac{1}{n} \sum_{t=1}^n \epsilon_t \epsilon_t' = \mathbf{I}_q + O_{\mathbb{P}}(n^{-1/2})$, so that assumption **C5** implies, for n sufficiently large, $\mathcal{S}_1 = o_{\mathbb{P}}(\sqrt{m})$. As for \mathcal{S}_2 , by the proof of lemma 1(b2) of Shimotsu (2007),

$$\mathcal{S}_2 = \frac{m}{n} O_{\mathbb{P}}(m^{1/2} \log m) = o_{\mathbb{P}}(\sqrt{m}),$$

where the last equality follows by assumption **C4**. Therefore,

$$\frac{1}{\sqrt{m}} \sum_{j=1}^m \lambda_j \Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) I_{\epsilon}(\lambda_j) \overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' = o_{\mathbb{P}}(1). \quad (\text{A.32})$$

Combining (A.31) and (A.32), we obtain

$$\begin{aligned} \frac{1}{\sqrt{m}} \sum_{j=1}^m (\lambda_j - \pi) \operatorname{Im} \left[\Lambda_j(\mathbf{d}_0)^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] &= \\ &= -\frac{\pi}{\sqrt{m}} \sum_{j=1}^m \operatorname{Im} \left[\Lambda_j(\mathbf{d}_0)^{-1} A(\lambda_j) I_{\epsilon}(\lambda_j) \overline{A(\lambda_j)}' \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] + o_{\mathbb{P}}(1). \end{aligned}$$

Again, by ignoring the $o_{\mathbb{P}}(1)$ term in (A.30) and proceeding as in p.298 of Shimotsu (2007) \mathcal{R}_2 can be approximated as

$$\mathcal{R}_2 = \sum_{u=2}^n \epsilon_u' \sum_{v=1}^{u-1} \tilde{\xi}_{u-v} \epsilon_v + o_{\mathbb{P}}(1), \quad \text{with} \quad \tilde{\xi}_t := \frac{1}{2n\sqrt{m}} \sum_{j=1}^m \operatorname{Re}[\Omega_j - \Omega_j'] \sin(t\lambda_j). \quad (\text{A.33})$$

Let

$$Z_1 := 0 \quad \text{and} \quad Z_t := \epsilon_t' \sum_{u=1}^{t-1} (\xi_{t-u} + \tilde{\xi}_{t-u}) \epsilon_u, \quad \text{for } t = 2, 3, \dots, n.$$

Combining (A.29) and (A.33), we conclude that

$$\sum_{k=1}^q \eta_k \sqrt{m} \frac{\partial S(\mathbf{d})}{\partial d_k} \Big|_{\mathbf{d}_0} = \sum_{t=1}^n Z_t + o_{\mathbb{P}}(1).$$

Observing that $\{Z_t\}_{t=1}^{\infty}$ is a zero mean martingale difference, by a standard martingale difference CLT, (A.20) follows if (cf. Billingsley, 1995, theorem 35.12)

$$\sum_{t=1}^n \mathbb{E}(Z_t^2 | \mathcal{F}_{t-1}) - \sum_{k=1}^q \sum_{l=1}^q \eta_k \eta_l \Omega_{kl} \xrightarrow{\mathbb{P}} 0, \quad \text{as } n \rightarrow \infty, \quad (\text{A.34})$$

and

$$\sum_{t=1}^n \mathbb{E}(Z_t^2 I(|Z_t| > \delta)) \xrightarrow{\mathbb{P}} 0, \quad \text{for all } \delta > 0, \text{ as } n \rightarrow \infty. \quad (\text{A.35})$$

Both results follow from the arguments in the proof of theorem 2, p.298, in Shimotsu (2007), since (A.34) and (A.35) are (29) and (30), respectively, in the aforementioned paper. Also observe that the $O_{\mathbb{P}}(1)$ terms in (A.25) and (A.30) do not affect the results, since, from our arguments,

$$\frac{1}{\sqrt{m}} \sum_{j=1}^m a_j \text{Re} \left[\Lambda_j(\mathbf{d}_0)^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right] \quad \text{and} \quad \frac{1}{\sqrt{m}} \sum_{j=1}^m \text{Im} \left[\Lambda_j(\mathbf{d}_0)^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' \right]$$

are both $O_{\mathbb{P}}(1)$. This completes the proof of (A.20).

We now move to show (A.21). For fixed $\delta > 0$, let $\boldsymbol{\theta} := \mathbf{d} - \mathbf{d}_0$ and define

$$\mathcal{M} := \left\{ \mathbf{d} : \log(n)^4 \|\mathbf{d} - \mathbf{d}_0\|_{\infty} < \delta \right\} = \left\{ \boldsymbol{\theta} : \log(n)^4 \|\boldsymbol{\theta}\|_{\infty} < \delta \right\}.$$

First we show that $\mathbb{P}(\bar{\mathbf{d}} \in \mathcal{M}) \rightarrow 1$, as $n \rightarrow \infty$. Assuming the same notation as in the proof of Theorem 4.2, recall the decomposition of $L(\mathbf{d}) = S(\mathbf{d}) - S(\mathbf{d}_0)$ given in expression (A.2). By applying the same argument as in the proof of Theorem 4.2, we first obtain

$$\inf_{\Theta_1 \setminus \mathcal{M}} \{\mathcal{R}(\mathbf{d})\} \geq \delta^2 \log(n)^8,$$

and upon applying Lemma 5.1, we obtain

$$\sup_{\Theta_1} \left\{ |\mathcal{A}(\mathbf{d}) - h(\mathbf{d})| \right\} = O_{\mathbb{P}} \left(\frac{m^{\alpha}}{n^{\alpha}} + \frac{\log(m)}{m^{2\epsilon}} + \frac{m}{n} \right).$$

It follows, uniformly in Θ_1 (cf. Shimotsu, 2007, p.300), that

$$\begin{aligned} \log(\mathcal{A}(\mathbf{d})) - \log(\mathcal{B}(\mathbf{d})) &\geq \log(h(\mathbf{d}) + o_{\mathbb{P}}(\log(n)^{-8})) - \log(h(\mathbf{d})) = o_{\mathbb{P}}(\log(n)^{-8}) \\ \log(\mathcal{A}(\mathbf{d}_0)) - \log(\mathcal{B}(\mathbf{d}_0)) &= \log(h(\mathbf{d}_0) + o_{\mathbb{P}}(\log(n)^{-8})) - \log(h(\mathbf{d}_0)) = o_{\mathbb{P}}(\log(n)^{-8}). \end{aligned}$$

Hence, $\mathbb{P}(\inf_{\Theta_1 \setminus \mathcal{M}} L(\mathbf{d}) \leq 0) \rightarrow 0$ and $\mathbb{P}(\bar{\mathbf{d}} \in \mathcal{M}) \rightarrow 1$, as $n \rightarrow \infty$, follows.

Now, observe that

$$\frac{\partial^2 S(\mathbf{d})}{\partial d_r \partial d_s} = \text{tr} \left[-\widehat{G}(\mathbf{d})^{-1} \frac{\partial \widehat{G}(\mathbf{d})}{\partial d_r} \widehat{G}(\mathbf{d})^{-1} \frac{\partial \widehat{G}(\mathbf{d})}{\partial d_s} + \widehat{G}(\mathbf{d})^{-1} \frac{\partial^2 \widehat{G}(\mathbf{d})}{\partial d_r \partial d_s} \right].$$

The derivatives of $\widehat{G}(\mathbf{d})$ are given by

$$\frac{\partial \widehat{G}(\mathbf{d})}{\partial d_r} = \frac{1}{m} \sum_{j=1}^m \text{Re} \left[\varphi(\lambda_j) \mathbf{I}_{(r)} \Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] + \frac{1}{m} \sum_{j=1}^m \text{Re} \left[\overline{\varphi(\lambda_j)} \Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \mathbf{I}_{(r)} \right], \quad (\text{A.36})$$

and

$$\begin{aligned} \frac{\partial^2 \widehat{G}(\mathbf{d})}{\partial d_r \partial d_s} &= \frac{1}{m} \sum_{j=1}^m \text{Re} \left[\varphi(\lambda_j)^2 \mathbf{I}_{(r)} \mathbf{I}_{(s)} \Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] + \\ &\quad + \frac{1}{m} \sum_{j=1}^m \text{Re} \left[|\overline{\varphi(\lambda_j)}|^2 \mathbf{I}_{(r)} \Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \mathbf{I}_{(s)} \right] + \\ &\quad + \frac{1}{m} \sum_{j=1}^m \text{Re} \left[|\overline{\varphi(\lambda_j)}|^2 \mathbf{I}_{(s)} \Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \mathbf{I}_{(r)} \right] + \\ &\quad + \frac{1}{m} \sum_{j=1}^m \text{Re} \left[\overline{\varphi(\lambda_j)}^2 \Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \mathbf{I}_{(r)} \mathbf{I}_{(s)} \right], \end{aligned}$$

where φ is given by (A.22). For $k = 0, 1, 2$, let

$$\begin{aligned}\mathcal{R}_k(\mathbf{d}) &= \frac{1}{m} \sum_{j=1}^m \log(\lambda_j)^k \operatorname{Re} \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] \\ \mathcal{I}_k(\mathbf{d}) &= \frac{1}{m} \sum_{j=1}^m \log(\lambda_j)^k \operatorname{Im} \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right],\end{aligned}$$

so that, we can write

$$\frac{\partial \widehat{G}(\mathbf{d})}{\partial d_r} = \mathbf{I}_{(r)} \mathcal{R}_1(\mathbf{d}) + \mathcal{R}_1(\mathbf{d}) \mathbf{I}_{(r)} + \frac{\pi}{2} \left(\mathbf{I}_{(r)} \mathcal{I}_0(\mathbf{d}) - \mathcal{I}_0(\mathbf{d}) \mathbf{I}_{(r)} \right) + o_{\mathbb{P}} \left(\frac{1}{\log(n)} \right), \quad (\text{A.37})$$

and

$$\begin{aligned}\frac{\partial^2 \widehat{G}(\mathbf{d})}{\partial d_r \partial d_s} &= \frac{\pi^2}{4} \left[\mathbf{I}_{(r)} \mathbf{I}_{(s)} \mathcal{R}_0(\mathbf{d}) + \mathbf{I}_{(r)} \mathcal{R}_0(\mathbf{d}) \mathbf{I}_{(s)} + \mathbf{I}_{(s)} \mathcal{R}_0(\mathbf{d}) \mathbf{I}_{(r)} + \mathcal{R}_0(\mathbf{d}) \mathbf{I}_{(s)} \mathbf{I}_{(r)} \right] + \\ &\quad + \pi \left[\mathbf{I}_{(r)} \mathbf{I}_{(s)} \mathcal{I}_1(\mathbf{d}) + \mathcal{I}_1(\mathbf{d}) \mathbf{I}_{(r)} \mathbf{I}_{(s)} \right] + \mathbf{I}_{(r)} \mathbf{I}_{(s)} \mathcal{R}_2(\mathbf{d}) + \mathbf{I}_{(r)} \mathcal{R}_2(\mathbf{d}) \mathbf{I}_{(s)} + \\ &\quad + \mathbf{I}_{(s)} \mathcal{R}_2(\mathbf{d}) \mathbf{I}_{(r)} + \mathcal{R}_2(\mathbf{d}) \mathbf{I}_{(s)} \mathbf{I}_{(r)} + o_{\mathbb{P}}(1). \quad (\text{A.38})\end{aligned}$$

The order of the remainder term in (A.37) is obtained as follows. Rewrite the first term on the RHS of (A.36) as

$$\begin{aligned}&\frac{1}{m} \sum_{j=1}^m \operatorname{Re} \left[\varphi(\lambda_j) \mathbf{I}_{(r)} \left(\operatorname{Re} \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] + i \operatorname{Im} \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] \right) \right] = \\ &= \frac{1}{m} \sum_{j=1}^m \operatorname{Re} \left[\log(\lambda_j) \mathbf{I}_{(r)} \operatorname{Re} \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] + i \log(\lambda_j) \mathbf{I}_{(r)} \operatorname{Im} \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] + \right. \\ &\quad \left. + i \left(\frac{\lambda_j - \pi}{2} \right) \mathbf{I}_{(r)} \operatorname{Re} \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] - \left(\frac{\lambda_j - \pi}{2} \right) \mathbf{I}_{(r)} \operatorname{Im} \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] \right] \\ &= \frac{1}{m} \sum_{j=1}^m \left[\log(\lambda_j) \mathbf{I}_{(r)} \operatorname{Re} \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] - \left(\frac{\lambda_j - \pi}{2} \right) \mathbf{I}_{(r)} \operatorname{Im} \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] \right].\end{aligned}$$

By summation by parts, Lemma 5.1 and assumption **C4**

$$\begin{aligned}&\frac{1}{m} \sum_{j=1}^m \lambda_j \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' \right] \leq \frac{1}{m} \sum_{j=1}^{m-1} |\lambda_j - \lambda_{j+1}| \left\| \sum_{k=1}^j \Lambda_k(\mathbf{d})^{-1} f_n(\lambda_k) \overline{\Lambda_k(\mathbf{d})^{-1}}' \right\|_{\infty} + \\ &\quad + \frac{\lambda_m}{m} \left\| \sum_{k=1}^j \Lambda_k(\mathbf{d})^{-1} f_n(\lambda_k) \overline{\Lambda_k(\mathbf{d})^{-1}}' \right\|_{\infty} \\ &\leq \frac{1}{m} \frac{m-1}{n} \left[O_{\mathbb{P}} \left((m-1)^{1/2} \log(m-1) + \frac{(m-1)^{\alpha+1}}{n^{\alpha}} \right) + O \left(\frac{1}{m} \right) \right] + \\ &\quad + O \left(\frac{1}{n} \right) \left[O_{\mathbb{P}} \left(m^{1/2} \log(m) + \frac{m^{\alpha+1}}{n^{\alpha}} \right) + O \left(\frac{1}{m} \right) \right] \\ &= O_{\mathbb{P}} \left(\frac{m^{1/2} \log(m)}{n} + \frac{m^{\alpha}}{n^{\alpha}} \right) + O \left(\frac{1}{mn} \right) = o_{\mathbb{P}} \left(\frac{1}{\log(n)} \right),\end{aligned}$$

where the last equality follows from

$$\frac{m^{1/2} \log(m)}{n} \log(n) = \frac{m^{1/2}}{n^{1/2}} \frac{\log(m)}{n^{1/4}} \frac{\log(n)}{n^{1/4}} \longrightarrow 0,$$

as n tends to infinity, by assumption **C4**. The other term in (A.36) is dealt analogously. The remainder term in (A.38) involves

$$\frac{1}{m} \sum_{j=1}^m \rho(\lambda_j) \Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}',$$

for $\rho(\lambda_j)$ proportional to λ_j , λ_j^2 and $\lambda_j \log(\lambda_j)$. The order of the terms proportional to λ_j has already been obtained, while the one proportional to λ_j^2 is dealt analogously since $\lambda_j^2 = O(\lambda_j)$. The term proportional to $\lambda_j \log(\lambda_j)$ is $o_{\mathbb{P}}(1)$ since, by summation by parts, Lemma 5.1 and assumption **C4**,

$$\begin{aligned} & \left\| \frac{1}{m} \sum_{j=1}^m \lambda_j \log(\lambda_j) \left[\Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}} \right]' \right\|_{\infty} \leq \\ & \leq \frac{1}{m} \sum_{j=1}^{m-1} \left| \lambda_j \log(\lambda_j) - \lambda_{j+1} \log(\lambda_{j+1}) \right| \left\| \sum_{k=1}^j \Lambda_k(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_k(\mathbf{d})^{-1}} \right\|_{\infty}' + \\ & \quad + \frac{\lambda_m \log(\lambda_m)}{m} \left\| \sum_{j=1}^m \Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}} \right\|_{\infty}' \\ & = \frac{1}{m} O_{\mathbb{P}} \left((m-1)^{1/2} \log(m-1) + \frac{(m-1)^{\alpha+1}}{n^{\alpha}} \right) o(1) + o_{\mathbb{P}}(1) = o_{\mathbb{P}}(1). \end{aligned}$$

In order to finish the proof, it suffices to show that,

$$\mathcal{R}_k(\mathbf{d}) = G_0 \frac{1}{m} \sum_{j=1}^m \log(\lambda_j)^k + o_{\mathbb{P}}(\log(n)^{k-2}) \quad (\text{A.39})$$

and

$$\mathcal{I}_k(\mathbf{d}) = o_{\mathbb{P}}(\log(n)^{k-2}), \quad (\text{A.40})$$

uniformly in $\mathbf{d} \in \mathcal{M}$. Indeed, if (A.39) and (A.40) hold, upon defining for a matrix M ,

$$T_1(M, r) := \mathbf{I}_{(r)} M + M \mathbf{I}_{(r)}, \quad T_2(M, r, s) := \mathbf{I}_{(r)} \mathbf{I}_{(s)} M + \mathbf{I}_{(r)} M \mathbf{I}_{(s)} + \mathbf{I}_{(s)} M \mathbf{I}_{(r)} + M \mathbf{I}_{(r)} \mathbf{I}_{(s)}$$

and

$$T_3(M, r, s) := -\mathbf{I}_{(r)} \mathbf{I}_{(s)} M + \mathbf{I}_{(r)} M \mathbf{I}_{(s)} + \mathbf{I}_{(s)} M \mathbf{I}_{(r)} - M \mathbf{I}_{(r)} \mathbf{I}_{(s)},$$

it follows that (cf. Shimotsu, 2007, p.301)

$$\begin{aligned} \widehat{G}(\bar{\mathbf{d}})^{-1} &= G_0^{-1} + o_{\mathbb{P}}(\log(n)^{-2}), \quad \frac{\partial \widehat{G}(\bar{\mathbf{d}})}{\partial d_r} = \frac{1}{m} \sum_{j=1}^m \log(\lambda_j) T_1(G_0, r) + o_{\mathbb{P}}(\log(n)^{-1}), \\ \text{and} \quad \frac{\partial^2 \widehat{G}(\bar{\mathbf{d}})}{\partial d_r \partial d_s} &= \frac{1}{m} \sum_{j=1}^m \log(\lambda_j)^2 T_2(G_0, r, s) + \frac{\pi^2}{4} T_3(G_0, r, s) + o_{\mathbb{P}}(1). \end{aligned}$$

Since $\text{tr} \left[G_0^{-1} T_1(G_0, r) G_0^{-1} T_1(G_0, s) \right] = \text{tr} \left[G_0^{-1} T_2(G_0, r, s) \right]$ and

$$\frac{1}{m} \sum_{j=1}^m \log(\lambda_j)^2 - \left(\frac{1}{m} \sum_{j=1}^m \log(\lambda_j) \right)^2 \rightarrow 1,$$

we obtain

$$\frac{\partial^2 S(\mathbf{d})}{\partial d_r \partial d_s} = \text{tr} \left[G_0^{-1} T_2(G_0, r, s) + \frac{\pi^2}{4} G_0^{-1} T_3(G_0, r, s) \right] + o_{\mathbb{P}}(1),$$

from where (A.21) follows. We proceed to show (A.39) and (A.40). For $k = 0, 1, 2$, let

$$\mathcal{F}_k(\boldsymbol{\theta}) := \frac{1}{m} \sum_{j=1}^m \log(\lambda_j)^k \Lambda_j(\boldsymbol{\theta})^{-1} G_0 \overline{\Lambda_j(\boldsymbol{\theta})^{-1}}'.$$

Then, (A.39) and (A.40) follow if

$$\sup_{\mathbf{d} \in \mathcal{M}} \left\{ \left\| \frac{1}{m} \sum_{j=1}^m \log(\lambda_j)^k \Lambda_j(\mathbf{d})^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d})^{-1}}' - \mathcal{F}_k(\boldsymbol{\theta}) \right\|_{\infty} \right\} = o_{\mathbb{P}}(\log(n)^{k-2}), \quad (\text{A.41})$$

and

$$\sup_{\mathbf{d} \in \mathcal{M}} \left\{ \left\| \mathcal{F}_k(\boldsymbol{\theta}) - G_0 \frac{1}{m} \sum_{j=1}^m \log(\lambda_j)^k \right\|_{\infty} \right\} = o(\log(n)^{k-2}). \quad (\text{A.42})$$

Following Shimotsu (2007), p.302, notice that, by applying (A.3), we can rewrite (A.41) as

$$\sup_{\mathbf{d} \in \mathcal{M}} \left\{ \left\| \frac{1}{m} \sum_{j=1}^m \log(\lambda_j)^k \Lambda_j(\boldsymbol{\theta})^{-1} \left[\Lambda_j(\mathbf{d}_0)^{-1} f_n(\lambda_j) \overline{\Lambda_j(\mathbf{d}_0)^{-1}}' - G_0 \right] \overline{\Lambda_j(\boldsymbol{\theta})^{-1}}' \right\|_{\infty} \right\}.$$

Define $b_j(\boldsymbol{\theta}; k) := \log(\lambda_j)^k e^{i(\lambda_j - \pi)(\theta_r - \theta_s)/2} \lambda_j^{\theta_r + \theta_s}$, for $k = 0, 1, 2$. Then, by omitting the supremum, the (r, s) -th element of (A.41) is equal to

$$\begin{aligned} & \left| \frac{1}{m} \sum_{j=1}^m b_j(\boldsymbol{\theta}; k) \left[e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} f_n^{rs}(\lambda_j) - G_0^{rs} \right] \right| \leq \\ & \leq \frac{1}{m} \sum_{j=1}^{m-1} \left| b_j(\boldsymbol{\theta}; k) - b_{j+1}(\boldsymbol{\theta}; k) \right| \left| \sum_{l=1}^j e^{i(\lambda_l - \pi)(d_r^0 - d_s^0)/2} \lambda_l^{d_r^0 + d_s^0} f_n^{rs}(\lambda_l) - G_0^{rs} \right| \\ & \quad + \frac{b_m(\boldsymbol{\theta}; k)}{m} \left| \sum_{j=1}^m e^{i(\lambda_j - \pi)(d_r^0 - d_s^0)/2} \lambda_j^{d_r^0 + d_s^0} f_n^{rs}(\lambda_j) - G_0^{rs} \right|, \end{aligned} \quad (\text{A.43})$$

where the inequality follows from summation by parts. Now, since

$$b_j(\boldsymbol{\theta}; k) - b_{j+1}(\boldsymbol{\theta}; k) = O\left(\frac{\log(n)^k}{j}\right) \quad \text{and} \quad b_m(\mathbf{d}; k) = O(\log(n)^k),$$

uniformly in $\boldsymbol{\theta} \in \mathcal{M}$, for any $k = 0, 1, 2$, it follows by Lemma 5.1 and Remark 5.1 that (A.43) can be rewritten as

$$\begin{aligned} & O\left(\frac{\log(n)^k}{m}\right) \frac{1}{m} O_{\mathbb{P}}\left(\frac{m^{\alpha+1}}{n^{\alpha}} + m^{1/2} \log(m)\right) + \\ & \quad + O(\log(n)^k) \frac{1}{m} O_{\mathbb{P}}\left(\frac{m^{\alpha+1}}{n^{\alpha}} + m^{1/2} \log(m)\right) = o_{\mathbb{P}}(\log(n)^{k-2}), \end{aligned}$$

where the last equality follows from assumption C4 (see also Remark 5.1), because

$$\begin{aligned} \log(n)^2 \frac{1}{m} O_{\mathbb{P}}\left(\frac{m^{\alpha+1}}{n^{\alpha}} + m^{1/2} \log(m)\right) &= \left[\frac{\log(n)^2}{m^{1/2} \log(m)} + \frac{\log(m)}{m^{1/2}} \frac{\log(n)^2}{m^{1/2}} \right] O_{\mathbb{P}}(1) \\ &= o(1) O_{\mathbb{P}}(1) = o_{\mathbb{P}}(1). \end{aligned}$$

The other term is dealt analogously, so that (A.41) follows. As for (A.42), the result follows from the proof of theorem 2, p.302, in Shimotsu (2007) (notice that it does not depend on f_n). This completes the proof. \blacksquare

References

- [1] Billingsley, P. (1995). *Probability and Measure*. 3rd ed., New York: John Wiley.
- [2] Chiriac, R. and Voev, V. (2011). “Modeling and Forecasting Multivariate Realized Volatility”. *Journal of Applied Econometrics*, **26**(6), 922-947.
- [3] Fox, R. and Taqqu, M.S. (1986). “Large-Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series”. *The Annals of Statistics*, **14**(2), 517-532.
- [4] Giraitis, L. and Surgailis, D. (1990). “A Central Limit Theorem for Quadratic Forms in Strongly Dependent Linear Variables and its Application to Asymptotic Normality of Whittle’s Estimate”. *Probability Theory and Related Fields*, **86**, 87-104.

- [5] Grenander, U. (1951). "On Empirical Spectral Analysis of Stochastic Process". *Arkiv för Matematik*, **1**, 197-277.
- [6] Grenander, U. and Rosenblatt, M. (1953). "Statistical Spectral Analysis of Time Series Arising From Stationary Stochastic Processes". *Annals of Mathematical Statistics*, **24**, 537-558.
- [7] Heyde, C.C. and Senata, E. (1972). "Estimation Theory for Growth and Immigration Rates in a Multiplicative Process". *Journal of Applied Probability*, **9**, 235-256.
- [8] Künsch, H. (1987). "Statistical Aspects of Self-Similar Processes". In Prokhorov, Yu. and Sazanov, V.V. (Eds.), *Proceedings of the First World Congress of the Bernoulli Society*. Utrecht: VNU Science Press, 67-74.
- [9] Lobato, I.N. (1999). "A Semiparametric Two-Step Estimator in a Multivariate Long Memory Model". *Journal of Econometrics*, **90**, 129-153.
- [10] Luceño, A. (1996). "A Fast Likelihood Approximation for Vector General Linear Processes with Long Series: Application to Fractional Differencing". *Biometrika*, **83**(3), 603-614.
- [11] Nielsen, F.S. (2011). "Local Whittle Estimation of Multi-Variate Fractionally Integrated Processes". *Journal of Time Series Analysis*, **32**(3), 317-335.
- [12] Priestley, M.B. (1981). *Spectral Analysis and Time Series*. London: Academic Press.
- [13] Robinson, P.M. (1995a). "Log-Periodogram Regression of Time Series with Long Range Dependence". *Annals of Statistics*, **23**(3), 1048-1072.
- [14] Robinson, P.M. (1995b). "Gaussian Semiparametric Estimation of Long Range Dependence". *Annals of Statistics*, **23**(5), 1630-1661.
- [15] Shimotsu, K. (2007). "Gaussian Semiparametric Estimation of Multivariate Fractionally Integrated Processes". *Journal of Econometrics*, **137**, 277-310.
- [16] Shimotsu, K. and Phillips, P.C.B. (2005). "Exact Local Whittle Estimation of Fractional Integration". *Annals of Statistics*, **33**(4), 1890-1933.
- [17] Sowell, F. (1989). "Maximum Likelihood Estimation of Fractionally Integrated Time Series Models". Working Paper, Carnegie-Mellon University.
- [18] Tsay, W.-J. (2010). "Maximum Likelihood Estimation of Stationary Multivariate ARFIMA Processes". *Journal of Statistical Computation and Simulation*, **80**(7-8), 729-745.
- [19] Zygmund, A. (2002). *Trigonometric Series. Vol. I, II*. Cambridge: Cambridge University Press.